AD_________________

Award Number: W81XWH-11-2-0158

TITLE: Medical Robotic and Tele-surgical Simulation Education Research

PRINCIPAL INVESTIGATOR: Roger D. Smith, PhD

CONTRACTING ORGANIZATION: Adventist Health Systems Inc.
Orlando, FL 32803

REPORT DATE: September 2015

TYPE OF REPORT: ANNUAL

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
September 2015	Annual	1 Sep 2014 - 31 Aug 2015

4. TITLE AND SUBTITLE
Medical Robotic and Tele-surgical Simulation Education Research

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-2-0158

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Roger D. Smith, PhD

Email: roger.smith@flhosp.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
.
Adventist Health Systems Inc.
Orlando, FL 32803

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project is broken into two focus areas: telesurgery and surgical rehearsal. In each we are exploring various applications and extensions of the existing robotic surgical systems. Under telesurgery we are exploring the ability to perform telesurgery using a robot both across a state-wide and a nation-wide area based on the currently available technology. Under surgical rehearsal exploring designs for simulator systems which can be used to improve training and education of surgeons pursuing expertise in the use of robotic surgical systems. The focus is on unique forms of robotics which have not previously been addressed by simulation technologies.

15. SUBJECT TERMS-

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU		**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Introduction

This project is broken into two focus areas: telesurgery and surgical rehearsal. In each we are exploring various applications and extensions of the existing robotic surgical systems. Under telesurgery we are exploring the ability to perform telesurgery using a robot both across a state-wide and a nation-wide area based on the currently available technology. Under surgical rehearsal exploring designs for simulator systems which can be used to improve training and education of surgeons pursuing expertise in the use of robotic surgical systems. The focus is on unique forms of robotics which have not previously been addressed by simulation technologies.

Simulator Performance is an experiment that was part of the scope of the original project which began in 2011. We have received permission to complete that study during the time of this extension, but using funds remaining from the original project. Therefore, that project is included in this report.

The scope described in the statement of work in this report is otherwise limited to the activities approved under a funding extension which was provided in September 2014 and extends through August 2016. Reports on the original body of work are covered in previous annual reports.

Statement of Work

Telesurgery: Metropolitan Latency. Perform robotic surgical experiments between multiple campuses within a metropolitan area, between campuses across a state area, and across nationwide campuses.

Period 1 Milestone: Telesurgery state-wide latency data and report. Award + 360 days.
Period 2 Milestone: Telesurgery nationwide latency data and report. Award + 700 days.

Surgical Rehearsal. Develop virtual reality environment for training operating room staff in robotic surgery. Develop design for simulators in hard-tissue robotic surgery (spinal and orthopedic).

Period 1 Milestone: Spinal simulator design document. Award + 300 days.
Period 2 Milestone: OR team training virtual world environment. Award + 360 days.
Period 2 Milestone: Orthopedic surgery rehearsal validation report. Award + 720 days.

Project Management

Progress Summary.

- We have received Alpha and Pre-Beta builds of the robotic OR virtual world. Working with sub-awardee on revisions and bug fixes.
- The data collection for Simulator Performance is completed. We are performing data analysis now.
- We have developed a user analysis for the spinal robotic simulator.
- Telesurgery experiments with Denver are complete. Exploring connection test to Seattle rather than Los Angeles.

Schedule.
This schedule shows our expected progress and completion of the experiments remaining on this grant.

		Y4Q4			**Y5Q1**			**Y5Q2**		
		2015			**2015**				**2016**	
Category	**Project**	**June**	**July**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**	**Apr**	**May**
Simulator Evaluations										
	Simulator Performance									
Simulator Design										
	Robotic OR Virtual World									
	Spinal Robotics									
	Orthopedic Robotics									
Telesurgery										
	Comms Latency									

Budget.
Financial spending on the project is under the projected budget. We have added a new MD Fellow to the team. We have also used three summer internals who have made significant contributions to the progress of the research.

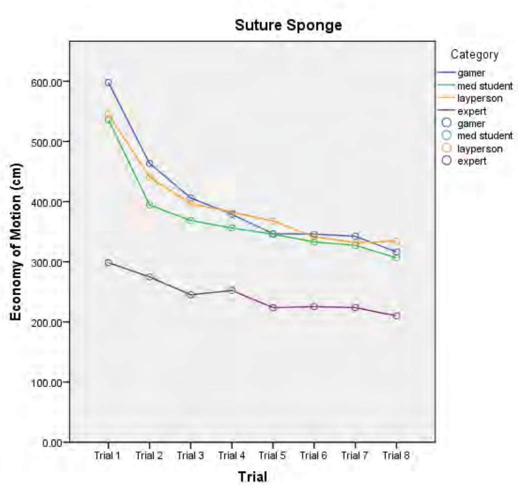
Scientific Progress

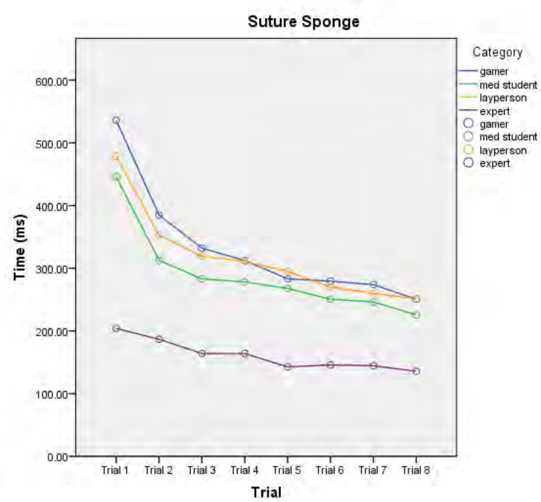
<u>*Simulator Performance.*</u>

Human subject data collection is finished. Expert surgeons were the most challenging population to collect. We collected data from Celebration Health, Columbia University Medical Center, and at the annual meeting of the Society for Laparoscopic Surgeons.

We are beginning full analysis of the data from all populations included in the study. The results of this work have been accepted for the 2015 I/ITSEC conference in December; serve as the basis for one doctoral dissertation; and are being prepared for journal submission.

A basic plot of the performance of four different populations while performing the Suture Sponge exercise is provide below (lower scores indicate higher skill levels in “economy of motion” and “time to complete”). This shows a very distinct performance difference between the expert surgeons and all other populations. It also appears that there is little difference between the population of lay people, medical students, and video gamers.

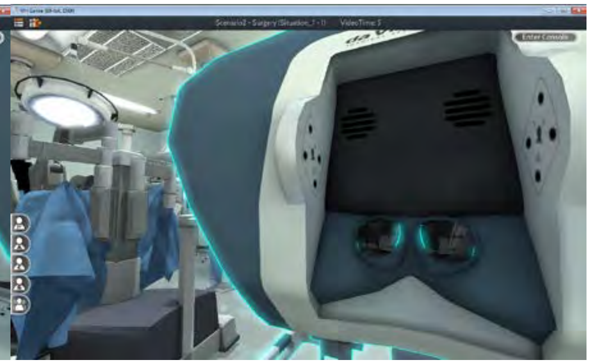




<u>*Simulator Design.*</u>

da Vinci OR Virtual World. Through meetings with expert robotic surgeons and their OR staff we have arrived at a design for a virtual world which may help surgeons to improve their team leadership skills in the OR. Working with ARA/Virtual Heroes Inc. we have created Alpha and Pre-Beta builds of a potential virtual world for OR leadership training based on TeamSTEPPS principles. The virtual world (or game) will be playable by a surgeon with all other roles being played by intelligent avatars. The avatars react to both correct and incorrect actions by the surgeon and provide spoken guidance toward the best behavior choices. Screen shots of the current build are provided below.

A paper on this project has been accepted at the 2015 I/ITSEC conference.



Spinal Robotic Simulator. We are using our knowledge of the Mazor Renaissance spinal robotic system to begin to understand how a simulator could be used to improve training for surgeons learning the system. Pieces of the surgeon's activities may be represented in: software on a laptop, hardware with electronic functionality, and inexpensive replicas of hardware.

During this quarter we have produced an analysis of the user requirements for such a simulator in a training environment. This is the precursor to a first design for such a device.

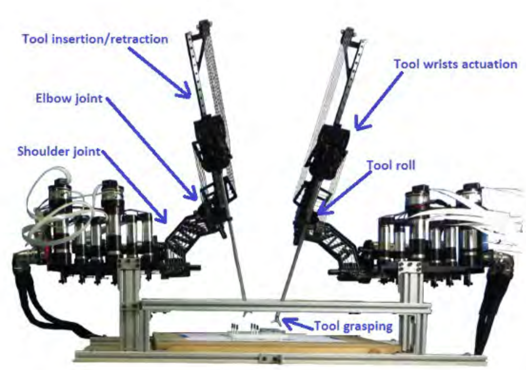
Orthopedic Robotic Simulator. This project will begin in 2016.

<u>*Telesurgery.*</u>

Orlando-to-Denver. Communication latency data collection is completed. Average two-way data transfer was 72 milliseconds. This speed is much faster than we had expected based on our inter-Florida experiments.

Centura Health Campus, Denver, CO

Los Angeles vs. Seattle. Original plans were to perform a telesurgery experiment from Orlando-to-Los Angeles. The necessary collaboration with that hospital has not materialized. Therefore we are now attempting to shift that experiment to a connection test with the University of Washington in Seattle. We are also expanding it to include data collection on the performance of the Raven II device.



Key Research Accomplishments

- *Telesurgery: Communications Latency*. Major hospital systems have sufficient telecommunication bandwidth to perform robotic telesurgery right now.
- *Surgical Rehearsal.* Simulation-based training for different forms of robotic procedures appears to be feasible beyond the simulators of the da Vinci robot which have previously been created. We are experimenting with (1) a virtual world for robotic OR team training, (2) a simulator to support training on the Mazor Renaissance spinal robotic device, and (3) a simulator to support orthopedic robotic procedures on knees and hips.
- Multiple publication and presentation have been generated from this research work.

Reportable Outcomes

Publications

Mouraviev et al. (Under Review). "Robotic training with porcine models induces less workload than virtual reality robotic simulators for urology resident trainees" Submitted to *Journal of the AUA* and AUA Annual Congress.

Smith, Tanaka, McIllwain, Willson. (Dec 2015). "Developing Game-based Leadership Training for Robotic Surgeons." *2015 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*.

Tanaka, Graddy, Smith, Perez. (Dec 2015). "Gamers Today, Surgeons Tomorrow?" *2015 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*.

Tanaka, Graddy, Simpson, Perez, Truong, & Smith. (Accepted). "Robotic Surgery Simulation Validity and Usability Comparative Analysis". *Journal of Surgical Endoscopy*.

Tanaka, Perez, Truong, & Smith. "From Design to Conception: An Assessment Device for Robotic Surgeons", *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*. December 2014. *Best Paper Nominee*

Tanaka, Graddy, & Smith. " Comparison of the Usability of Robotic Surgery Simulators", *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*. December 2014. *Honorable Mention for Best Paper*

Smith & Simpson. " Return on Investment for Robotic Surgical Simulators", *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*. December 2014. *Honorable Mention for Best Paper*

Smith, Truong, & Perez. (2014) Comparative analysis of the functionality of simulators of the da Vinci surgical robot. *J Surg Endosc*, 1-12.

Perez, Xu, Chauhan, Tanaka, Simpson, Abdul-Muhsin, & Smith. "Impact of delay on telesurgical performance: Study on the dV-Trainer robotic simulator". Submission to Journal of Urology 2014.

Smith, "The Future of Robotic Technology", Robotic Surgery of the Head and Neck, Springer Press, 2015 (projected).

Martino, Siddiqui, et al. "Fundamentals of Robotic Gynecologic Surgery" Developing a Quality Improvement Project to Improve Patient Safety", Society of Gynecologic Oncology, Annual Meeting on Women's Cancer, March 2014.

Smith, Patel, & Satava. "Fundamentals of robotic surgery: a course of basic robotic surgery skills based upon a 14-society consensus template of outcomes measures and curriculum development", *The International Journal of Medical Robotics and Computer Assisted Surgery*, October 2013. DOI: 10.1002/rcs.1559

Smith, "From FLS to FRS: The Fundamentals of Robotic Surgery are on their Way", World Robotic Gynecologic Congress, Chicago, IL. 2013

Advincula & Smith. "Contributions of Laparoscopic Surgical Experience to the Development of Robotic Proficiency", *Society for Gynecologic Surgery Annual Meeting*, March 2013.

Smith, Chauhan, & Satava. "Fundamentals of Robotic Surgery Consensus: Outcomes Measures and Curriculum Development", *NextMed: Medicine Meets Virtual Reality Conference*. February 2013.

Smith & Truong. "Robotic Surgical Education with Virtual Simulators", *2013 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*. December 2013.

Smith & Chauhan. "Using Simulators to Measure Communication Latency Effects in Robotic Telesurgery", *2012 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*. December 2012 *Best Paper Nominee*

Satava, Smith & Patel. "Report on the First Consensus Conference on the Fundamentals of Robotic Surgery" Outcomes Measures", *ACS Accredited Education Institutes Meeting*. March 2012.

Presentations

2015

Smith (August 2015). "The Validation of Surgical Simulators for RASD". FDA Workshop on Robotically Assisted Surgical Devices.

Tanaka, Graddy, Perez, Simpson, Truong, Smith. (Nov 2015). "Video Game Impact on Basic Robotic Surgical Skills." Annual Meeting of the Association of Gynecologic Laparoscopists.

Perez, Tanaka, Simpson, Truong, Smith, Satava. (Nov 2015). "From concept to surgical relevance: Engineering the training device for the Fundamentals of Robotic Surgery." Annual Meeting of the Association of Gynecologic Laparoscopists.

Tanaka, Perez, Graddy, & Smith. " Video Game Experience and Basic Robotic Skills", Florida Hospital Internal Research Forum, Orlando, FL, April 2015.

Truong, Tanaka, Simpson, Perez, & Smith. "Robotic surgical simulation versus traditional didactics for surgical training: a randomized controlled trial", Society for Gynecologic Surgeons Annual Scientific Meeting, Orlando, FL, March 2015.

Smith. "Update on Robotic Surgical Simulation", *2015 Society of Robotic Surgeons (SRS),* Orlando, FL, February 2015.

Smith. "Fundamentals of Robotic Surgery", *2015 Society of Robotic Surgeons (SRS),* Orlando, FL, February 2015.

2014

Truong, Tanaka, Simpson, Advincula, & Smith. "A Prospective Randomized Controlled Comparative Study on Surgical Training Methods and Impact on Surgical Performance: Virtual Reality Robotic Simulation vs. Didactic Lectures", AAGL Global Congress on Minimally Invasive Gynecology, November 2014

Smith & Simpson. "Return on Investment for Robotic Surgical Simulators", AAGL Global Congress on Minimally Invasive Gynecology, November 2014

Tanaka, Truong, & Smith. "Robotic Surgical Simulators: An Assessment of Usability and Preferences", AAGL Global Congress on Minimally Invasive Gynecology, November 2014

Simpson, Perez, Tanaka, Truong & Smith. “Validating the Efficacy of GEARS through the Assessment of 100 Videos”, Society of Laparoendoscopic Surgeons Annual Meeting & Endo Expo, September 2014.

Truong, Tanaka, Simpson, Perez, Smith & Advincula. “Randomized Controlled Study Comparing Robotic Simulation Versus Didactic Teaching for Robotic Surgical Training: Opinions and Perspectives”, Society of Laparoendoscopic Surgeons Annual Meeting & Endo Expo, September 2014. *Honorable Mention for the Paul Alan Wetter Award for Best MultiSpecialty Scientific Paper*

Smith & Tanaka. “Gamers in Surgical Simulation: A Comparison of Gamers, Surgeons, and Clinical Staff”, Defense GameTech Users Conference, Orlando, FL, September 2014.

Lendvay, Simpson, Truong, & Smith. “Differentiating Surgical Skill through the Wisdom of Crowds”, European Endoscopic Urology Society, April 2014.

Patel, Patel & Smith, “Feasibility of Robotic Telesurgery across a Multi-Campus Metropolitan Hospital System”, Third Biennial Miami Robotics Symposium, April 2014.

Smith, “Robotic & Telesurgery Research”, Stetson University Senior Tech Expo, March, 2014.

Satava & Smith, “Fundamentals of Robotic Surgery: Development and Validation of an Online Curriculum and New Psychomotor Testing Device”, NextMed/MMVR Conference, February, 2014.

Satava & Smith, “Fundamentals of Robotic Surgery: Development and Validation of an Online Curriculum and New Psychomotor Testing Device”, CAMLS-Halldale Summit on New Technology in Medicine, February, 2014.

Tanaka, Truong, Simpson, Perez, & Smith, “A Comparison of the Effectiveness and Usability of Robotic Simulators”, Florida Hospital Internal Research Forum, January 2014.

Smith, “Robotic Surgery Education, Simulation & Telesurgery”, Adventist Health System, Surgeon Executives Meeting, January 2014.

2013

Truong: "The Fundamentals of Robotic Surgery Psychomotor Skills Prototype Development Video": Harrith M Hasson Award for Best Presentation Promoting Education and Training, 2013 SLS Annual Meeting in Reston, Virginia. Smith, “Robotic Surgery Education, Simulation & Telesurgery, Society for Laparoscopic Surgeons, Fellowship Summit, December 2013.

Smith, “Virtual Reality Simulation: The Future”, Society for Robotic Surgery, Annual Meeting, November, 2013.

Smith, “Strategic Technology Leadership: The Role of the Technology Executive”, MITRE Leadership Forum, October 2013.

Smith, “Robots in the Hands of your Surgeon”, IEEE Orlando Chapter Annual Meeting, October 2013.

Smith, “Medical Simulation in Robotic Surgery”, Lou Frey Institute of Politics and Government, University of Central Florida, September 2013.

Smith, “Robots in the Hands of Your Surgeon”, Chinese American Scholars and Professionals Association of Florida, Miami Annual Meeting, August 2013.

Smith, “Innovation for Trainers”, *Training 2013 Conference*. Keynote Presentation, February 2013.

2012

Smith, “Simulation Surgeon, Soldier Spy”, Keynote presentation, *2012 SpringSim MultiConference*. March 2012.

Smith, “Robotic Surgery and Surgical Simulation”, presentation to *International Council on Systems Engineering – Orlando Chapter*. February 2012.

Smith, “Beyond Education and Training: Challenges of Running Medical Simulators in New Paradigms”. *2012 International Meeting on Simulation in Healthcare*. January 2012.

2011

Smith, "Simulation in Surgical Education", American College of Healthcare Executives, December 2011.

Smith, "Medical Simulation Special Event: Robotic and Telesurgery Research Using Simulation", I/ITSEC, December 2011.

Smith, "Robotic and Telesurgery Research", National Center for Simulation, October 2011.

Smith, "Medical Simulation Standards: What can we learn from the DoD?" *Medical Technology, Training, and Treatment Conference*, May 2011.

Smith, "Simulation and Game Technology in Medical Education", IDEAS Workshop, Harvard Medical School, April 2011.

Smith, "Robotic Surgery and Surgical Simulation", Guest Lecture, Old Dominion University, April 2011.

2010

Smith, "Surgical Simulation Research Initiatives", I/ITSEC UCF Workshop, December 2010.

Smith, "da Vinci Surgical Robot", I/ITSEC techPATH Teachers Workshop, November 2010.

Poster Presentations

2015

Smith, Simpson. “Return on Investment Model for Robotic Simulators”, Poster Presentation at *2015 Society of Robotic Surgeons (SRS),* Orlando, FL, February 2015.

Tanaka, Graddy, Abdul-Muhsin, Simpson, Truong, & Smith. “A Comparison of Validity and Usability of Robotic Simulation”, Poster Presentation at *2015 Society of Robotic Surgeons (SRS)*, Orlando, FL, February 2015.

Tanaka, Perez, & Smith. “ Fundamentals of Robotic Surgery Psychomotor Skills: Metrics Development and Evaluation”, Poster Presentation at *2015 International Meeting on Simulation in Healthcare (IMSH) Conference,* New Orleans, LA, January 2015.

2014

Lendvay TS, White LW, Holst D, Kowalewski T, Harper JD, Sorenson M, Brand TC, Truong M, Simpson K, Smith R. Quantifying Surgical Skill Using the Wisdom of Crowds. *American*

College of Surgeons Clinical Congress, San Francisco, CA, October 26-30th, 2014. [Poster #PP2014-51161].

Lendvay T, Holst D, White L, Kowalewski T, Brand T, Sorenson M, Harper J, Truong M, Simpson K, Smith R. Differentiating Surgical Skill Through the Wisdom of Crowds. *American Urological Association Annual Meeting*, Engineers in Urology Session, Orlando, FL, May 16-21, 2014 [Moderated Poster #82].

Awards

Nominee, 2015 Florida Hospital Des Cummings Innovator Award

Best Paper Nominee in the Emerging Concepts and Innovative Technologies track, *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference,* Orlando, FL.

Honorable Mention for Best Paper in the Training track *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*

Honorable Mention for Best Paper in the Policy, Standards, Management and Acquisition track *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*, Orlando, FL.

2014 Florida Hospital Des Cummings Innovator Award

Best Paper Nominee in the Human Performance track *2014 Interservice/Industry Training Education and Simulation (I/ITSEC) Conference*

2014 Honorable Mention for the Paul Alan Wetter Award for Best Multispecialty Scientific Paper, MIS Week Annual Conference

2013 Second Place, Top Gun Surgery Competition, MIS Week Annual Conference

2013 Harrith M. Hasson Award for Best Presentation Promoting Education and Training, MIS Week Annual Conference

2013 Best Video Session in Multispecialty Surgery, MIS Week Annual Conference

2013 Silver Medal, Robotic Surgery Olympics, MIS Week Annual Conference

2013 Third Place, Top Gun Surgery Competition, MIS Week Annual Conference

2012 Schwartz Industry Innovation Award, Orlando Economic Development Commission. The Nicholson Center’s research work in robotic telesurgery was recognized locally as one of the most innovative activities in the Orlando metropolitan area.

2012 Best Paper Award for Simulation Technologies Track, Interservice/Industry Training Education and Simulation Conference (I/ITSEC)

Conclusion

Each of the research areas funded by this grant has made significant scientific contributions. The knowledge gained from this work is being shared through reports to the government and multiple presentations at both clinical and simulation conferences. We have also submitted multiple papers for journal publication.

This cooperative agreement is scheduled to end on August 31, 2016. Based on our current work flow and state of funds the project is currently on schedule to complete all objectives by the end of the agreement.

Appendices

Copies of manuscripts, abstracts, and presentations of work resulting from this grant are included as appendices to this report.

Robotic Surgery Simulation Validity and Usability Comparative Analysis

Journal:	*Surgical Endoscopy*
Manuscript ID:	SEND-15-0350
Manuscript Type:	Original Article
Date Submitted by the Author:	04-Mar-2015
Complete List of Authors:	Tanaka, Alyssa; Florida Hospital, Nicholson Center Graddy, Courtney; Florida Hospital Celebration Health, Simpson, Khara; Columbia University Medical Center, Perez, Manuela; University of Lorraine-Nancy, Truong, Mireille; Columbia University Medical Center, Smith, Roger; Florida Hospital, Nicholson Center
Keyword:	simulation, Validation, Usability, Robotic Surgery, Training , Education
Please specify the country from which you are submitting your manuscript.:	United States

Dr. Mark A. Talamini
Editor-In-Chief
March 5, 2015

Dear Editor,

Please find enclosed an article we wish to submit for publication in The Journal of Surgical Endoscopy entitled: "Robotic Surgery Simulation Validity and Usability Comparative Analysis." Our team has been conducting work relating to minimally invasive surgery for many years. We seek publication in The Journal of Surgical Endoscopy because it provides the surgical community a juncture to exchange critical information on practice, theory, and research in various medical and surgical disciplines. The evaluation of the usability and validity of available training tools is valuable and relevant information to in the surgical community. Also, this paper details the second phase of work previous published in the journal entitled "Comparative analysis of the functionality of simulators of the da Vinci surgical robot."

This manuscript has not been published and is not being considered for publication elsewhere. There are no financial or other relations that could lead to a conflict of interest. Each author has contributed significantly to the submitted work:
1) Conception and Design: Roger Smith, Mirelle Truong.
2) Data acquisition: Alyssa Tanaka, Courtney Graddy
3) Data Analysis and interpretation: Alyssa Tanaka, Khara Simpson, Courtney Graddy
4) Drafting the manuscript: Alyssa Tanaka, Courtney Graddy, Khara Simpson, Manuela Perez
5) Critical Review of the Manuscript: all authors
6) Supervision: Roger Smith

All authors have read and approved the final version of this manuscript.
The address for correspondence is:

Alyssa Tanaka,
404 Celebration Pl.
Celebration, FL 34747
Phone: (407) 303-4276; Mobile: (321) 480-5510; E-mail: Alyssa.tanaka@flhosp.org

We appreciate your time and consideration.

Respectfully,
The authors.

Robotic Surgery Simulation Validity and Usability Comparative Analysis

Running Head: Robotic simulation validity and usability

Authors:

Alyssa Tanaka, MS, Florida Hospital Nicholson Center, Celebration, Florida
Courtney Graddy, MHA, Florida Hospital Celebration Health, Celebration, Florida
Khara Simpson, MD, Columbia University Medical Center, New York, NY
Manuela Perez, MD, PhD, University of Lorraine-Nancy, Nancy, FR
Mireille Truong, MD, Columbia University Medical Center, New York, NY
Roger Smith, PhD, Florida Hospital Nicholson Center, Celebration, Florida

Corresponding Author:
Alyssa Tanaka, Florida Hospital Nicholson Center, 404 Celebration Place, Celebration, FL 34747, e-mail: Alyssa.tanaka@flhosp.org
(v) 407-303-4276
(f) 407-303-4473

Funding Source: U.S. Army Telemedicine and Advanced Technology Research Center. Grant #: W81XWH-11-2-0158

ABSTRACT

Background: The introduction of simulation into minimally invasive robotic surgery is relatively recent and has seen rapid advancement; therefore, a need exists to develop training curriculums and identify systems that will be most effective at training surgical skills. Several simulators have been introduced to support these aims -- the daVinci Skills Simulator, Mimic dV-Trainer, Surgical Simulated Systems' RoSS, and Simbionix Robotix Mentor. While multiple studies have been conducted to demonstrate the validity of these systems, studies comparing the perceived value of these devices as tools for education and skills are lacking.

Methods: Subjects who qualified as medical students or physicians (n=105) were assigned a specific order to use each of the three simulators. After completing a demographic questionnaire, participants performed one exercise on the three simulators and completed a second questionnaire regarding their experience with the device. After using all systems, they completed a final questionnaire, which detailed their comparative preferences. The subject's performance metrics were also collected from each simulator.

Results: The data confirmed the face, content, and construct validity for the dV-Trainer and Skills Simulator. Similar validities could not be confirmed for the RoSS. Greater than 80% of the time, participants chose the Skills Simulator in terms of physical comfort, ergonomics, and overall choice. However, only 55% thought the skills simulator was worth the cost of the equipment. The dV-Trainer had the highest cost preference scores with 71% percent of respondents feeling it was worth the investment.

Conclusions: Usability can affect the consistency and commitment of users of robotic surgical simulators. In a previous study, these simulators were objectively reviewed and compared in terms of their system capabilities. Collectively, this work will offer end users and potential buyers a comparison of the perceived value and preferences of robotic simulators.

KEYWORDS

Simulation; Validation; Robotic Surgery; Training; Usability

INTRODUCTION

Medicine has come to the conclusion that the Halstedian training model (i.e., See one, do one, teach one) is no longer sufficient for teaching complex skills, particularly robotic surgical skills [1]. With the introduction of robotic technology between patient and surgeon, a need to master new skills has emerged. A number of virtual reality simulators have been developed to support the training and acquisition of such skills. Currently, the commercially available robotic simulators include: the da Vinci Skills Simulator (dVSS) by Intuitive Surgical Inc., also known as the "Backpack Simulator"; the dV-Trainer from Mimic Technologies Inc.; the RoSS by Simulated Surgical Sciences LLC; and the Robotix Mentor from Simbionix (Figure 1). All of these da Vinci simulators utilize a visual scene that is presented in a computer-generated 3D environment providing challenging tests for practicing dexterity and machine operations. Originally, the simulated exercises trained basic robotic skills; however with advances in technology, surgeons can now train for specific procedures (e.g. partial nephrectomy and hysterectomy).

Figure 1. Simulators of the da Vinci robotic surgical system

The work described in this paper is the second part of a three-phase analysis to study the effectiveness of these simulators and applications to the education of robotic surgeons. In the first phase, the authors evaluated and compared the objective characteristics of three simulators (dVSS, dV-Trainer, and RoSS). The Simbionix Robotix Mentor was not included because it was under development at the time of this research. This analysis provided a head-to-head comparison of the systems and found that they varied greatly in their hardware and software.

In the dVSS, the trainee operates the simulated environment using the actual da Vinci surgical console. The simulator is a custom computer, appended to the surgical console through the surgical data port. While the simulator costs approximately $85,000, the surgical console costs $500,000 incurring an investment of $585,000. Using this simulator, users can train with the actual hardware they would use during surgery; however, this requires availability of the surgical console, which may be fully scheduled in the operating room. Few hospitals have a dedicated training console, meaning that users do not have ready access to the simulator. The second system is a standalone system that utilizes a high performance graphic/gaming computer, connected to a custom desktop viewing and control device that replicates the hardware of the da Vinci surgeon's console. This system shares similar software with the dVSS, but does not require the use of actual da Vinci hardware. The cost of this simulator is approximately $96,000. The third system is composed of a completely customized replica of the da Vinci surgeon's console. Internally the simulator contains a graphic computer, a 3D viewing system, and commercial Omni Phantom haptic controllers. This simulator uses unique software and costs approximately $126,000 [2].

The validity of medical and surgical simulators is typically evaluated using the categories defined by McDougal [3]. This paper defines the most commonly recognized forms of validation as: *face, content, construct, concurrent, and predictive validity*. *Face validity* is typically assessed informally by users and indicates whether the simulator is an accurate representation of the actual system (i.e. the realism of the simulator). *Content validity* is the measure of the appropriateness of the system as a teaching modality. Experts who are knowledgeable about the device typically assess this via a formal evaluation. *Construct validity* is the ability of a simulator to differentiate between the performances of experienced users and those who are novices. *Concurrent validity* is the extent to which the simulator correlates with the "gold standard" for training and *predictive validity* is the extent to which the simulator can predict a user's future surgical performance. Collectively, concurrent and predictive validity are known as criterion validity and are used as measures of the simulator's ability to correlate trainee performance with their real life performance. Face and content validity are most effective in evaluating the ability of a simulator to train a surgeon; however construct, concurrent, and predictive validity are most useful for evaluating the effectiveness of a simulator to assess a trainee.

The validity of all three simulators has been examined separately (Table 1) and to our knowledge there is no comparative research of all three systems. The current study therefore compares the three commercially available da Vinci simulators and details the findings for face, content, and construct validity of these systems. The purpose of this is to provide end-users and potential buyers with a head-to-head evaluation of the value and usability of the systems.

Table 1. da Vinci simulator validation studies from Smith R, Truong M, & Perez M [2]

MATERIALS AND METHODS

Participants in this study included medical students, residents, fellows, and attending physicians. Participants were recruited from the University of Central Florida College of Medicine, courses held at the Florida Hospital Nicholson Center, and two surgical robotics conferences (World Robotics Gynecology Congress and Society of Robotic Surgeons Scientific Meeting). Subjects were excluded from participating if they had participated in a formal robotic simulation-training course to eliminate preference bias. Each participant was categorized into one of three groups (i.e. Expert, Intermediate, or Novice) according to the self-reported number of robotic cases performed. Individuals who had performed 0-19 robotic cases, were categorized as Novices, individuals with 20-99 robotic cases were considered to be Intermediates, and individuals with 100 or more cases were considered to be Experts.

After being categorized into an experience level, each participant was assigned a specific order in which they used each of the simulators (Figure 2). This alternating order was implemented to identify and eliminate any potential bias that may exist by using a specific system first. All participants completed one exercise on each of the simulators. The tasks chosen were Peg Board 1 in both the dV-Trainer and the dVSS and Ball Placement 1 in the RoSS. The same task was used for both the dV-Trainer and the dVSS because these systems share similar software and exercises. The RoSS software contains unique exercises and Ball Placement 1 was chosen because it trains the same basic skills as Peg Board 1.

Figure 2. Example of rotating order and research process

After completing the exercise on a simulator, participants completed a post-questionnaire (Survey 1), which asked for feedback regarding their experience on that specific simulator. After using all three systems, subjects completed a second post-questionnaire (Survey 2), which asked them to compare all three systems to each other. The participant's performance metrics were also collected from each of the simulators.

RESULTS

The Novice group (n=37) had performed an average of 2 robotic cases, the Intermediate group (n=31) on average performed 54 cases, and the Expert group (n=37) performed 336 cases. Sixty-two percent of subjects were men and 38% were women with an average age of 43. On average, participants had 15 years in practice and 3 years of robotic experience. Seventy-six percent were attending physicians and 73% of participants were currently or had received robotic surgery training, while 41% provided that they train residents and fellows. A one-way ANOVA verified a difference in the average age and number of years in practice of participants based on the classification of expert, intermediate or novice (number of robotic procedures). This is to be expected since higher ages typically imply a higher number of years of practice and resultant larger numbers of robotic procedures.

The types of validity evaluated in this experiment were face, content, and construct. To analyze the systems for face validity and content validity, questions from Survey 1 were used. The questions were evaluated on a five point Likert scale (i.e., Strongly Disagree, Disagree, Neither Agree or Disagree, Agree, and Strongly Agree). As recommended by Van Nortwick et al. [26], face validity was analyzed by expert and intermediate feedback only as these are the users most familiar with the robotic system; however, only expert feedback was used for content validity because they have the best ability to judge the appropriateness of the system as a training tool. For construct validity, performance metrics such as Overall Score, Time to Complete, Number of Errors, and Economy of Motion were analyzed (Table 2). Specifically, Time and Economy of Motion were chosen due to a previous study by Perrenot, Perez, Tran, Jehl, Felblinger, Bresler, & Hubert [10] indicating that these are highly relevant indicators of expertise in robotic surgery.

Table 2. Description of data used for types of validity.

Face Validity

A Chi-square test of independence was used to evaluate the distribution of scores for a specific simulator in relation to the order of the system's presentation to the subject. This analysis indicated that there was no difference in participants' responses according to the order in which the systems were presented; and established that no bias was present due to the presentation order ($p>0.05$). These questions asked participants to evaluate whether the hand controllers on the simulator were effective for working in the simulated environment (Question 1) and if the device is a sufficiently accurate representation of the real robotic system (Question 4). For both questions, the RoSS had the lowest average score, dV-Trainer had the second highest score, and the dVSS had the highest score of the three (Table 3). A repeated measures ANOVA verified that the answers were statistically different for both questions ($p<0.001$).

Table 3. Mean scores from a 5-point Likert scale on face validity

Content Validity

As seen in Table 4, 100% of participants either agreed or strongly agreed that the 3D graphical exercises in the dVSS were effective for teaching robotic skills while 59% disagreed or strongly disagreed that the RoSS' capabilities were effective. When asked if the scoring system effectively communicated their performance, 88% of dVSS users agreed or strongly agreed, while 79% of dV-Trainer users agreed or strongly agreed. Similarly, 91% and 82% of participants agreed or strongly agreed that the dVSS and dV-Trainer, respectively, effectively guided them to improve their performance, while only 36% felt the RoSS provided the same guidance.

Table 4. Percentages of Likert responses for content validity questions

Construct Validity

The overall score, number of errors, time to complete, and economy of motion scores collected by the simulators for Experts (n=37) and Novices (n=37) were used to compare construct validity (Table 5). Intermediate subjects were not included in the construct validity analysis because it was only necessary to determine if the simulator could distinguish specifically between novice and expert users. Overall Score is synthesized from multiple metrics and is specific to the individual simulator. This metric was available in the dVSS and the dV-Trainer, however the Overall Score metric is not automatically exported by the RoSS and therefore was not analyzed for this system. Instead, the Number of Errors was used for the RoSS. For all of the simulators, higher Overall Score values are better, while lower Economy of Motion, Time, and Number of Error values are better preferred.

For the RoSS, the analysis has 23 missing data points because the system does not report scores when a user exceeds a maximum exercise time or chooses to terminate the exercise before completion. This resulted in a sample of 30 experts and 21 novices on this system. A Mann-Whitney U test showed that the distributions of time ($p=0.221$), number of errors ($p=0.644$), and economy of motion ($p=0.566$) were not statistically different for the experts compared to the novice group on this simulator.

The dV-Trainer analysis of experts (n=37) and novices (n=37) had three missing values for economy of motion and completion time and five for the overall score metric, thus the analysis contained varying number of subjects. The distribution of the overall scores was not significantly different for the expert compared to the novice group ($p=0.061$). These tests did confirm statistical differences for economy of motion ($p<0.001$) and time to complete ($p<0.001$), with a lower economy of motion value and shorter completion time for experts compared to novices.

The dVSS analysis included all novice (n=37) and expert (n=37) participants. Time to complete ($p<0.001$) and overall score ($p=0.006$) were significantly different for the expert compared to the novice group. The expert group had a higher overall score and a shorter completion time compared to the novice group. However, economy of motion did not show a statistical difference with this analysis ($p=0.216$).

Table 5. Mann-Whitney U test level of significance on construct validity measures

The relationship between experience and performance metrics was more specifically analyzed in terms of the self-reported number of cases of all participants (n=105) using a non-parametric correlation coefficient (Spearman's). For the RoSS, 30 participants were excluded from the analysis. For the participants that were included in the analysis (n=75), there was not a significant correlation between Time to Complete ($p=0.181$), Number of Errors ($p=0.563$), or Economy of Motion ($p=0.390$) with the total number of robotic cases performed (Figure 3).

Figure 3. Graphs of correlation between experience and metrics on the RoSS

For the dV-Trainer, four participants were excluded from the entire analysis and two participants were excluded from the Overall Score analysis (Overall Score n=99; Economy of Motion and Time to Complete n=101). The analysis verified a statistically significant correlation between Overall Score ($p=0.03$), Economy of Motion ($p<0.01$), and Time to Complete ($p<0.01$). The correlation value was negative for Economy of Motion and Time to Complete, showing that with a greater number of robotic cases, the time taken and distance moved decreased. The correlation was positive for Overall Score indicating that the participants' score increased with the number of robotic cases performed (Figure 4).

Figure 4. Graphs of correlation between experience and metrics on the dV-Trainer

For the dVSS, two participants were excluded from the analysis (n=103). A statistically significant difference was found between Overall Score ($p=0.01$) and Time to Complete ($p<0.01$). The correlation value was negative for time and positive for Overall Score, signifying that with more robotic cases the time taken decreased and the score increased. There was not a statistically significant correlation between Economy of Motion and the total number of robotic cases performed ($p=0.105$) (Figure 5).

Figure 5. Graphs of correlation between experience and metrics on the dVSS

Usability (Preference)

The questions from Survey 2 were used to understand the preference of the subjects when using the simulators. All subjects were included in this analysis except for two participants who were dropped from the analysis because they did not complete the questionnaire. The participants' responses to the following usability questions can be seen in Figure 6:

- *If you are (were) a program director, which simulator would you choose for your trainees;*
- *In which simulator were you physically more comfortable;*
- *Which simulator had the best hand controls;*
- *Which simulator had the best foot controls;*
- *Which simulator had the best 3D vision;*
- *Were you feeling stressed or annoyed by any of the simulators?*

Figure 6. Description of usability responses

Overall, most participants preferred the dVSS and indicated that they would choose this device as a training system if they were a program director. Participants not only felt most comfortable in the dVSS, but also felt that the system had the best control and vision equipment. The least preferred system was the RoSS, which most participants also agreed made them feel stressed or annoyed. Ten percent of participants also responded that they felt stressed or annoyed by both the dV-Trainer (dVT) and the RoSS.

Cost

All participants were also asked to provide feedback on their simulator preference in terms of the cost of the system. The responses were analyzed in terms of the frequency of the responses given. Most participants felt that the dV-Trainer was worth the investment; while most felt that the RoSS was not. When asked about the dVSS, only 56% of participants agreed that it was worth the investment (Figure 7).

Figure 7. Description of cost preferences

DISCUSSION

The aim of this study was to conduct a comparison of the three commercially available simulators used to train surgeons on the daVinci robotic surgical system. The study was performed to assist potential buyers in making a purchasing and deployment decision regarding robotic simulators. This study provides information about the face, content, and construct validity, as well as usability of the systems.

The simulators were perceived to be different in their representation of the real robotic system. The dVSS was most preferred in terms of ergonomics and usability; however, most participants did not feel that this system was worth the investment. The costs provided in the questionnaire included all equipment needed to make the simulator functional. While the simulator itself only costs $85,000, it is impossible to use without the $500,000 da Vinci surgeon console. By leveraging the actual da Vinci hardware, this simulator allows for a more realistic experience, but limits the availability and creates a higher cost for training than other robotic simulators. Economy of Motion was not able to differentiate novices from experts in the dVSS, which could be attributed to the ease of use of the controllers allowing novices to move the controls as efficiently as experts. The generous workspace of the dVSS could also have an impact on the lack of difference.

In terms of cost, most participants agreed that the dV-Trainer had the best cost-effectiveness. In contrast to the dVSS, the dV-Trainer is a standalone simulator and does not require the support of the daVinci hardware to operate. This allows for better accessibility and requires less of an investment for training. The Overall Score aspect of construct validity in the dV-Trainer may not have shown a difference between novices and experts due to the way that the scoring is developed. The scoring system is constructed with a “ceiling” that prevents users from achieving a high Overall Score without attaining high scores across multiple metrics.

The RoSS was the least preferred system for comfort and other usability aspects (i.e., hand controls, foot controls, and 3D interface), with most participants feeling stressed or annoyed when using the system. This study was unable to validate the face, content, or construct validity for this system. Currently, there is limited data available that confirms construct validity of the RoSS. Similarly to Raza [21], this study was unable to confirm a difference between experts and novices in terms of time taken to complete the exercise. As stated previously, time and economy of motion are considered highly relevant measures of expertise levels [10] and should distinguish between these groups in the simulators.

To our knowledge this three-part study is the first to compare three of the available simulators. This study involved the largest sample size and diversity of participants (i.e., experience levels, number of robotic cases, and subspecialty type) thus far in relevant publications. The results from this research will help guide the choice of simulators used for future studies at Florida Hospital and may also influence decisions at other laboratories. However, a limitation to the study was the lack of consistency in the available exercises and scoring systems across the three systems. A consideration for future studies will be to use more complex exercises and increase the depth of the face and content validity evaluation. Future research is also necessary to evaluate and compare new iterations of da Vinci simulators (e.g the Simbionix Robotix Mentor).

ACKNOWLEDGEMENTS

The authors wish to thank the representatives of each simulator company for their assistance in collecting images.

da Vinci Skills Simulator photos ©2013 Intuitive Surgical, Inc. Used with permission.
dV-Trainer Simulator photos ©2013 Mimic Technologies, Inc. Used with permission.
RoSS Simulator photos ©2013 Simulated Surgical Systems, LLC. Used with permission.
Robotix Mentor photos ©2014 Simbionix USA, Corp. Used with permission.

DISCLOSURES

Alyssa Tanaka has no conflicts of interest or financial ties to disclose.
Courtney Graddy has no conflicts of interest or financial ties to disclose.
Dr. Khara Simpson has no conflicts of interest or financial ties to disclose.
Dr. Manuela Perez has no conflicts of interest or financial ties to disclose.
Dr. Mireille Truong has no conflicts of interest or financial ties to disclose.
Dr. Roger Smith has no conflicts of interest or financial ties to disclose.

Source of Funding: U.S. Army Telemedicine and Advanced Technology Research Center. Grant #: W81XWH-11-2-0158

REFERENCES

1. Cameron JL (1997) William Stewart Halsted. Our surgical heritage. Annals of surgery 225(5):445

2. Smith R, Truong M, & Perez M (2014) Comparative analysis of the functionality of simulators of the da Vinci surgical robot. J Surg endosc 1-12

3. McDougall EM (2007) Validation of surgical simulators. J Endourol 21(3):244-7

4. Hung AJ, Zehnder P, Patil MB (2011) Face, content and construct validity of a novel robotic surgery simulator. J Urology 186(3):1019-24.

5. Kelly DC, Margules AC, Kundavaram CR (2012) Face, content, and construct validation of the da Vinci Skills Simulator. J Urology 79(5):1068-72

6. Liss MA, Abdelshehid C, Quach S (2012) Validation, Correlation, and Comparison of the da Vinci Trainer(™) and the da Vinci Surgical Skills Simulator(™) Using the Mimic(™) Software for Urologic Robotic Surgical Education. J Endourol 26(12):1629-34

7. Lendvay, Casale P, Sweet R, Peters C (2008) Initial validation of a virtual-reality robotic simulator. J of Rob Surg 2(3):145-149

8. Kenney P, Wszolek M, Gould J, Libertino J, Moinzadeh A (2009) Face, Content, and Construct Validity of dV-Trainer, a Novel Virtual Reality Simulator for Robotic Surgery. J Urology 73(6):1288-92

9. Sethi AS, Peine WJ, Mohammadi Y (2009) Validation of a novel virtual reality robotic simulator. J Endourol 23(3):503-8

10. Perrenot C, Perez M, Tran N, Jehl JP, Felblinger J, Bresler L, & Hubert J (2012) The virtual reality simulator dV-Trainer(®) is a valid assessment tool for robotic surgical skills J Surg Endosc 26(9):2587-93

11. Korets R, Mues AC, Graversen JA, Gupta M, Benson MC, Cooper KL, Landman J, Badani KK (2011) Validating the Use of the Mimic dV-Trainer for Robotic Surgery Skill Acquisition Among Urology Residents. J Urology 78(6):1326-30

12. Lee JY, Mucksavage P, Kerbl DC (2012) Validation study of a virtual reality robotic simulator--role as an assessment tool? J Urology 187(3):998-1002

13. Schreuder HWR, Persson JEU, Wolswijk RGH, Ihse I, Schijven MP, Verheijen RHM (2014) Validation of a novel virtual reality simulator for robotic surgery. ScientificWorldJournal 2014:507076. doi: 10.1155/2014/507076

14. Seixas-Mikelus SA, Kesavadas T, Srimathveeravalli G (2010) Face validation of a novel robotic surgical simulator. J Urology 76(2):357-60

15. Stegemann AP, Kesavadas T, Rehman S, Sharif M, Rao A, DuPont N, Shi Y, Wilding,G, Hassett J, and Guru K (2012) Development, Implementation and Validation of a Simulation-Based Curriculum for Robot-Assisted Surgery. Presented at the AUA poster session, May 19-23, 2012 Atlanta, GA.

16. Hung AJ, Patil MB, Zehnder P (2012) Concurrent and predictive validation of a novel robotic surgery simulator: a prospective, randomized study. J Urology 187(2):630-7

17. Colaco M, Balica A, Su D (2012) Initial experiences with RoSS surgical simulator in residency training: a validity and model analysis. J Rob Surg 7(1):71-75

18. Finnegan KT, Meraney AM, Staff I (2012) da Vinci Skills Simulator construct validation study: correlation of prior robotic experience with overall score and time score simulator performance. J Urology 80(2):330-5

19. Connolly M, Seligman J, Kastenmeier A, Goldblatt M, Gould JC (2014) Validation of a virtual reality-based robotic surgical skills curriculum. Surg Endosc. doi: 10.1007/s00464-013-3373-x

20. Lendvay TS, Brand TC, White L, Kowalewski T, Jonnadula S, Mercer LD, Khorsand D, Andros J, Hannaford B, Satava RM (2013) Virtual reality robotic surgery warm-up improves task performance in a dry laboratory environment: a prospective randomized controlled study. J Am Coll Surg 216:1181–1192

21. Raza SJ, Chowriappa A, Stegemann A, Ahmed K, Shi Y, Wilding G, Kaouk J, Peabody J, Menon M, Kesavadas T, Guru KA. (2013). Construct Validation of the Fundamental Skills of Robotic Surgery (FSRS) Curriculum. Podium presentation, AUA Annual Congress, San Diego.

22. Tergas AI, Sheth SB, Green IC, Giuntoli RL 2nd, Winder AD, Fader AN (2013) A pilot study of surgical training using a virtual robotic surgery simulator. JSLS J Soc Laparoendosc Surg Soc Laparoendosc Surg 17:219–226

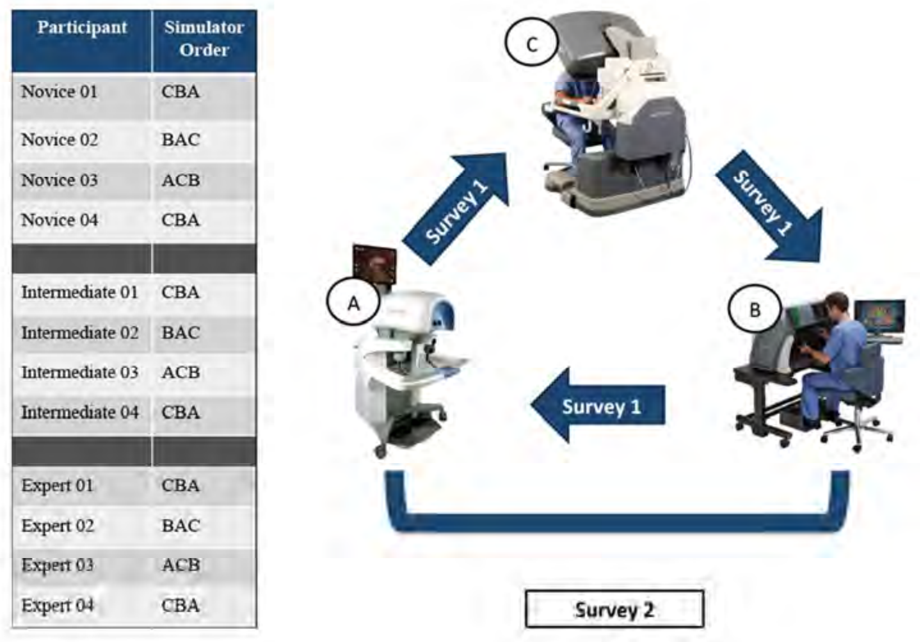
23. Lerner M, Ayalew M, Peine W, Sundaram C (2010) Does Training on a Virtual Reality Robotic Simulator Improve Performance on the *da Vinci*® Surgical System? J Endourol 24(3):467-72

24. Chowriappa A, Raza SJ, Stegemann A, Ahmed K, Shi Y, Wilding G, Kaouk J, Peabody JO, Menon M, Kesavadas T, Guru KA (2013) Development & Validation of a Composite Grading & Scoring System for Robot-assisted Surgical Training – The Robotic Skills Assessment Scale. Podium presentation, AUA Annual Congress, San Diego.

25. Culligan P, Gurshumov E, Lewis C, Priestley J, Komar J, Salamon C (2014) Predictive validity of a training protocol using a robotic surgery simulator. Female Pelvic Med Reconstr Surg 20:48–51

26. Van Nortwick S, Lendvay T, Jensen A, Wright A, Horvath K, & Kim S (2010) Methodologies for establishing validity in surgical simulation studies. Surgery 147 (5), 622-630

Figure 1. Simulators of the da Vinci robotic surgical system
76x50mm (300 x 300 DPI)



Participant	Simulator Order
Novice 01	CBA
Novice 02	BAC
Novice 03	ACB
Novice 04	CBA
Intermediate 01	CBA
Intermediate 02	BAC
Intermediate 03	ACB
Intermediate 04	CBA
Expert 01	CBA
Expert 02	BAC
Expert 03	ACB
Expert 04	CBA

Figure 2. Example of rotating order and research process
76x50mm (300 x 300 DPI)

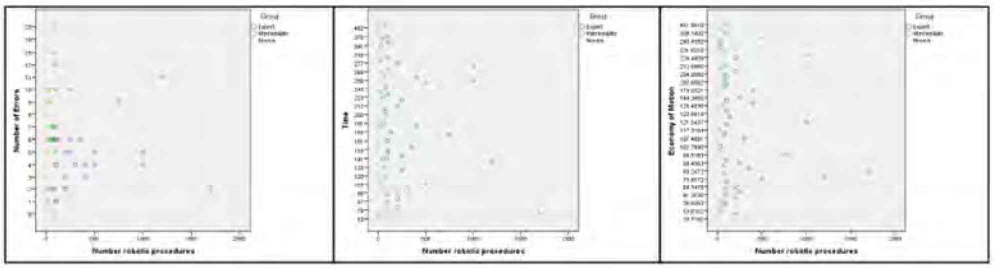


Figure 3. Graphs of correlation between experience and metrics on the RoSS
76x50mm (300 x 300 DPI)

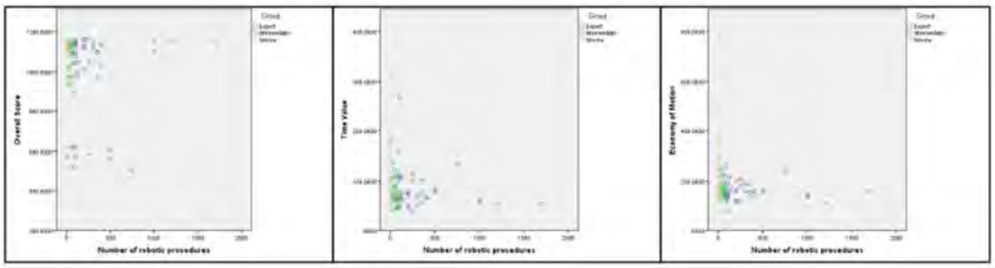


Figure 4. Graphs of correlation between experience and metrics on the dV-Trainer
76x50mm (300 x 300 DPI)

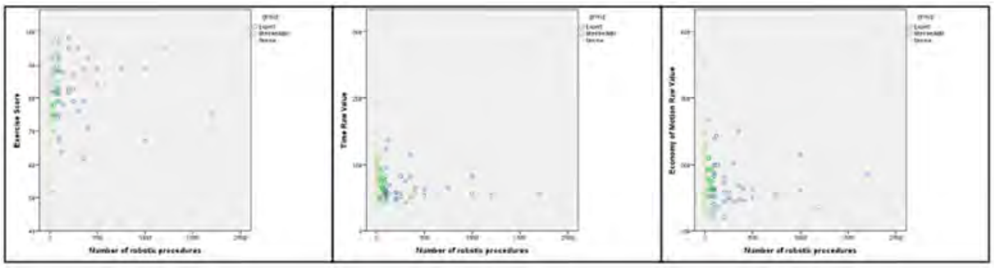


Figure 5. Graphs of correlation between experience and metrics on the dVSS
76x50mm (300 x 300 DPI)

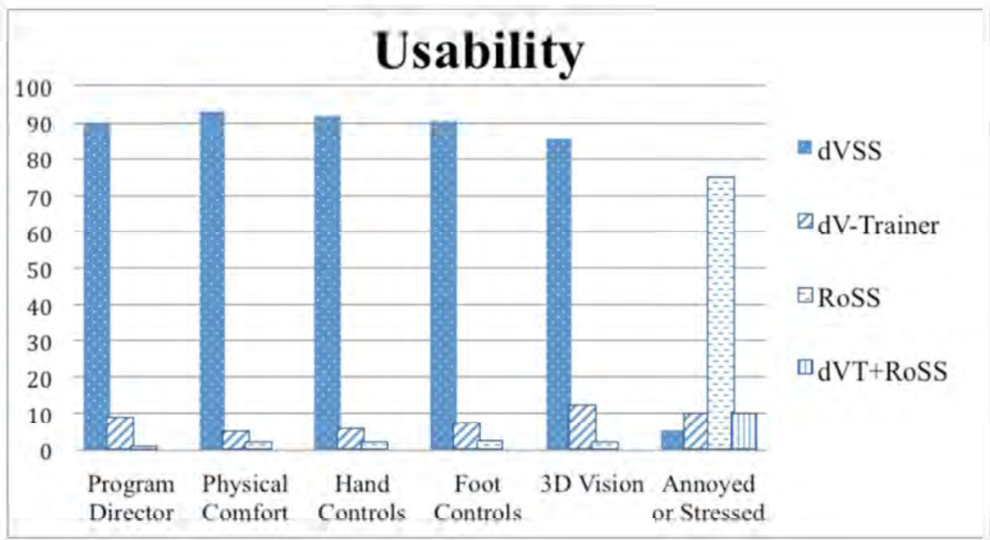


Figure 6. Description of usability responses
76x50mm (300 x 300 DPI)

Figure 7. Description of cost preferences
76x50mm (300 x 300 DPI)

Table 1. da Vinci simulator validation studies from Smith R, Truong M, & Perez M [2]

Validation	DVSS	dV-Trainer	RoSS
Face: ***Subjective realism of the simulator***	*Hung [4]* *Kelly [5]* *Liss [6]*	*Lendvay [7]* *Kenney [8]* *Sethi [9]* *Perrenot [10]* *Korets [11]* *Lee [12]* *Schreuder [13]*	*Seixas-Mikelus [14]* *Stegemann, [15]*
Content: ***Judgment of appropriateness as a teaching modality***	*Hung [4]* Hung *[20]* *Kelly [5]* *Liss [6]*	*Kenney [8]* *Sethi [9]* *Perrenot [10]* *Lee [12]*	*Seixas-Mikelus [14]* *Colaco [17]*
Construct: ***Able to distinguish experienced from inexperienced surgeon***	*Hung [4]* *Kelly [5]* *Liss [6]* *Finnegan [18]*	*Kenney [8]* *Perrenot [10]* *Korets [11]* *Lee [12]* *Schreuder [13]* *Connolly [19]* Lendvay *[20]*	*Raza [21]*
Concurrent: ***Extent to which simulator correlates with "gold standard"***	*Hung [16]* Tergas *[22]*	*Perrenot [10]* *Korets [11]* *Lee [12]* *Lerner [23]*	*Chowriappa, [24]*
Predictive: ***Extent to which simulator predicts future performance***	*Hung [16]* Tergas *[22]* Culligan *[25]*		

Table 2. Description of data used for types of validity

Type of Validity	Evaluation	Type of Participant	Question/Metric
Face Validity	Survey 1	Expert and Intermediate	Q1: The hand controllers on this simulator are effective for working in the simulated environment (Likert).
			Q4: The device is a sufficiently accurate representation of the real robotic system (Likert).
Content Validity	Survey 1	Expert	Q2: The 3D graphical exercises in the simulator are effective for teaching robotic skills (Likert).
			Q5: The scoring system effectively communicates my performance on the exercise (Likert).
			Q6: The scoring system effectively guides me to improve performance on the simulator (Likert).
Construct Validity	Simulator	Experts and Novices	Overall Score (points)
			Number of Errors (count)
			Time to Complete (seconds)
			Economy of Motion (centimeters)

Table 3. Mean scores from a 5-point Likert scale on face validity

Face Validity (n=68)	DVSS	dV-Trainer	RoSS
Q1: The hand controllers on this simulator are effective for working in the simulated environment.	4.80	3.62	2.17
Q4: The device is a sufficiently accurate representation of the real robotic system.	4.65	3.45	1.82

Table 4. Percentages of Likert responses for content validity questions

Content Validity (n=34)					
Likert Score	**Strong Dis**	**Disagree**	**Neither**	**Agree**	**Strong Agree**
Q2: *The 3D graphical exercises in the simulator are effective for teaching robotic skills.*					
DVSS	0%	0%	0%	35.3%	64.7%
dV-Trainer	2.9%	5.9%	11.8%	50.0%	29.4%
RoSS	20.6%	38.2%	17.6%	17.6%	5.9%
Q5: *The scoring system effectively communicates my performance on the exercise.*					
DVSS	2.9%	5.9%	2.9%	38.2%	50.0%
dV-Trainer	2.9%	2.9%	14.7%	55.9%	23.5%
RoSS	17.6%	20.6%	26.5%	29.4%	5.9%
Q6: *The scoring system effectively guides me to improve performance on the simulator.*					
DVSS	0%	0%	8.8%	61.8%	29.4%
dV-Trainer	2.9%	2.9%	11.8%	61.8%	20.6%
RoSS	18.2%	18.2%	27.3%	33.3%	3.0%

Table 5. Mann-Whitney U test level of significance on construct validity measures

	DVSS	dV-Trainer	RoSS
Time to Complete	p<0.001	p<0.001	p=0.221
Overall Score	p<0.01	p=0.061	n/a
Economy of Motion	p=0.216	p<0.001	p=0.566
Number of Errors	n/a	n/a	p=0.644

GEARS Study

SLS Abstract

Validating the Efficacy of GEARS through the Assessment of 100 Videos

Objective: To evaluate the use of the GEARS assessment to distinguish varying robotic surgical experience. To identify additional variables that may increase the reliability and accuracy of GEARS.

Methods: Videos were collected for 104 medical students and surgeons performing a simple cystotomy closure on an animate model. Subjects were divided into three groups based on their robotic experience. Reviews were performed by three surgeons (inter-rater reliability 0.95 cronbach's alpha) using GEARS, errors and tasks assessments. Statistical analysis was performed to determine the validity and internal consistency of GEARS as well as its correlation with the independently developed errors and tasks check lists.

Results: There were differences between task time, GEARS, error and task metrics across experience groups (p = 0.01). For the individual categories of GEARS, all were able to differentiate between experience groups. Bimanual dexterity and efficiency were the best at differentiating intermediates (26-100 robotic cases) from experts (>100). In the evaluation of errors, missed targets, needle drops, tissue damage, and instrument collisions were statistically different among groups and instruments out of view, needle re-loading, tissue re-grasping were not. Instrument collision and needle drops showed the greatest differentiation among groups. GEARS had internal consistency (cronbach's alpha 0.88) and construct validity (p=<0.001).

Conclusions: This is the first study to confirm the ability of GEARS to differentiate robotic surgical experience at all levels of training. While errors and task assessment also showed differences among experience levels, they were less reliable than GEARS.

INTRODUCTION:

The inception of robotic surgery has changed the face of surgery by providing an additional means to completing more complex procedures in a min. invasive fashion. While there are obvious parallels with traditional laparoscopic surgery, endowrist manipulation and lack of haptic feedback represent significant differences. Due to its unique interface, robotics has necessitated device specific training and evaluation. To this aim, several robotic surgical simulators have been developed and validated. The Fundamentals of Robotic Surgery (FRS) is also being trialed and validated to offer specialty non-specific didactic and simulation training, including a high stakes test.

For objective in-OR and video assessment, the global evaluative assessment of robotic skills (GEARS) is used most often. GEARS, similar to GOALS (a validated laparoscopic objective assessment tool), is a six-item objective assessment using a Likert scale. The six fundamental skills assessed are: depth perception, bimanual dexterity, tissue handling, efficiency, autonomy, and robotic control. The first five skills are the same as GOALS. The five point likert scale has definitions associated with 3 of the 5 scores. GEARS has been found to have construct validity, and good intra and inter-rater reliability (Goh et al 2012). No studies have been done evaluating its ability to differentiate levels of expertise and which skills sets/categories best correlate to proficiency/surgeon expertise level. We propose to evaluate the above as well as to test additional structured objective assessments (error based and task based scoring) to better categorize GEARS.

Our hypothesis is that collisions and tissue damage will be highly associated with novice performance and instrument out of view will be a less discriminatory factor for experts.

BACKGROUND

GOALS was created from the concept of OSATS which includes the objective assessment of skills and a task assessment.

METHODS

Archived videos and assessments were reviewed for 100 surgeons performing a simple cystotomy closure. Previous assessements included the GEARS and an addendum questionnaire regarding errors and task assessment. The errors were designed to correlate with each of the GEARS items. (Look below). Reviews were performed by three independent reviewers whose interrater reliability was .95 via cronbach's alpha. The trained reviewers were all familiar with robotic surgery and included two minimally invasive fellows and an attending surgeon. Training consisted of reviewing the GEARS, encouraging full use of the scale, and a focus on documentation of only confirmed errors. Reviewers were blinded to the experience level of the surgeons. Intra questionnaire correlation was performed within the GEARS as well as with the addendum questionnaire to determine internal consistency. An ICC was performed for the total GEARs score as well as with the deletion of each item. A similar process was performed for the task and error portion of the questionnaire. Finally the overall scores of each (GEARS, task, error) were compared to see if there was a relationship. The mean and median scores were calculated for each based on experience levels. Compare each item with experience level to determine which item most correlated with the three experience categories. Scatterplot graphs were performed comparing the number of robotic surgeries peformed with the total GEARS scores, errors, and tasks. Confirm that the developed errors and task questions correlated with the initially identified items in GEARS.

GENERAL ANALYSIS

1. Mean and median score for each experience group for the GEARS, errors, and task assessment. Determine if these differences are statistically different. HYPOTHESIS: there will not be a statistically significant difference between intermediates and experts but it will be present between novices and experts. RESULTS – There were statistically significant differences across all experience levels.

2. Identify which components of the GEARS or errors or task assessments where there are statistically significant differences – item by item by analysis.

3. Looking at correlation between the items with no statistical significance to determine if associated task or error components gave better differentiation between the two groups.

SUB-ANALYSIS

1. Identify relationships between the associations between GEARS items and the errors and task (see below) HYPOTHESIS:

2. Come up with a total score/equation including the GEARS, errors (neg scoring), and tasks….compare that to the different experience levels and see if there is a statistically significant difference among the groups

RESULTS

CONCLUSION/DISCUSSION

Considerations – We did not fully evaluate the benchmark of autonomy. As noted above subjects were given the same task to complete without assistance. Reviewers in most cases gave

LIT REVIEW: Google scholar search term GEARS

<u>GEARS: Original study</u>

Jan 2012 Goh et al – Journal of Urology

- Used expert consensus to identify an objective assessment tool for the eval of robotic surgeons – GOALS with an additional variable for robotic control

- Determined to have good internal consistency and reliability. Also established construct validity but interestingly enough could not differentiate PGY-6 (avg surgical cases 30) from expert level surgeons (avg surgical cases 190)

- 4 attg surgeons, 25 trainees (PGY4-6)

<u>Teaching surgical skills – Reznick NEJM 2006</u>

Reviews changes in medical education and considerations for the future

Traditional halstedian model of education stems from education through volume of patients. With current advances in medicine the people are living longer, increasing the complexity of conditions. Certain conditions are becoming increasingly more rare as well. We also have to contend with work hour shortages. To improve/strengthen training and exposure – we are turning to simulation for orientation and to improve outcomes.

Learning Tools and simulation in Robotic Surgery: State of the Art – Citation Pubmed

-Fitts and Posner's 3 stage theory of motor skill acquisition – suggests that in the first phase the learner is more focused on the steps of the material and typically is less coordinated. This suggests that maybe this stage of learning should be performed in a non-lethal environment

Perform searches on FLS/OSATS/Goals/Include RTN

Conclusions – (Discussion) Unable to differentiate between intermediate and expert level surgeons with GEARS alone. Need another variable or term for surgical skills.

-Limitations – Did not fully assess autonomy – everyone defaulted to a 5.

Meeting Notes 4/28/14

Correlations between GEARS and errors

1. Depth perception – Intstrumentation
2. Bimanual dexterity – Needle Handling and
3. Tissue Handling – Tissue handling
4. Efficiency – Time
5. Autonomy – Nothing/didn't fully evaluate
6. Robotic Skills – Instrumentation and needle handling and tasks

REVIEW

Current status of robotic simulators in acquisition of robotic surgical skills

Anup Kumar[a], Roger Smith[b], and Vipul R. Patel[a]

Purpose of review
This article provides an overview of the current status of simulator systems in robotic surgery training curriculum, focusing on available simulators for training, their comparison, new technologies introduced in simulation focusing on concepts of training along with existing challenges and future perspectives of simulator training in robotic surgery.

Recent findings
The different virtual reality simulators available in the market like dVSS, dVT, RoSS, ProMIS and SEP have shown face, content and construct validity in robotic skills training for novices outside the operating room. Recently, augmented reality simulators like HoST, Maestro AR and RobotiX Mentor have been introduced in robotic training providing a more realistic operating environment, emphasizing more on procedure-specific robotic training . Further, the Xperience Team Trainer, which provides training to console surgeon and bed-side assistant simultaneously, has been recently introduced to emphasize the importance of teamwork and proper coordination.

Summary
Simulator training holds an important place in current robotic training curriculum of future robotic surgeons. There is a need for more procedure-specific augmented reality simulator training, utilizing advancements in computing and graphical capabilities for new innovations in simulator technology. Further studies are required to establish its cost–benefit ratio along with concurrent and predictive validity.

Keywords
robotics surgery, simulation, surgical training, virtual reality

INTRODUCTION

The use of the robotic platform in urology has expanded exponentially over the last decade and has established itself in most advanced centres across the world, particularly in the USA [1–3]. In 2013, approximately 80% of radical prostatectomies were performed using robotic platform in the USA [1]. This tremendous growth in robotic technology has highlighted the increasing demand for surgeons trained in robotic skills. Although most urology residency programs are presently incorporating robotic surgery as a part of their curriculum, adequate training of these future robotic surgeons is facing many challenges [4–6]. First, there has been a decrease in actual training hours along with risk of litigation, increased emphasis on patient safety and improved surgical outcomes. Second, the traditional Halstedian method of training of 'see one, do one and teach one' does not apply to robotic technology. The robot-assisted radical prostatectomy is a complex procedure requiring complete knowledge of pelvic anatomy and an understanding of magnification, depth perception, three-dimensional spatial orientation and coordinated hand–eye movements. Third, in robotics, the mentor is not working close to the trainee with one person at the console and one other person required for bedside assistance, thus raising concerns in the mentor's mind about the patient's safety [7–9]. The training can be divided as preclinical and clinical [4–6]. The preclinical training includes use of simulators, defined as tools

[a]Department of Urology, Global Robotics Institute, Florida Hospital-Celebration Health and [b]Chief Technology Officer, Florida Hospital-Nicholson Center, University of Central Florida College of Medicine, Orlando, Florida, USA

Correspondence to Dr Anup Kumar, Department of Urology, Global Robotics Institute, Florida Hospital-Celebration Health and University of Central Florida College of Medicine, Orlando, FL 32827, USA. Tel: +1 407 701 0365; e-mail: anup_14k@yahoo.com

Curr Opin Urol 2015, 25:000–000

DOI:10.1097/MOU.0000000000000137

KEY POINTS

- The simulator training can form an integral part of credentialing and training robotic surgery of future robotic surgeons.
- It has the potential to decrease the learning curve for the acquisition of robotic skills.
- It can supplement the hands-on training clinical phase and can act as a bridge between preclinical training and actual hands-on clinical training without jeopardizing the safety of patients.
- There is a need for more procedure-specific augmented reality simulator training in a cost-effective manner, with more emphasis on both technical skills and team-work training.

enabling the operator to reproduce or represent under test conditions a phenomenon likely to occur in actual performance. Clinical training includes observation, bed-side assistance and hands-on-training under mentorship (including Tele-mentoring) and proctoring [4,7,10].

Simulators can be classified as low fidelity, high fidelity, virtual reality and augmented reality [4–7,11,12■■]. Low fidelity simulators, like Dry lab laparoscopic box trainer, are portable, less expensive and have been proven to improve surgical skills over time. But, they have disadvantages of lack of duplication of a real surgical environment, lack of feedback and inability to teach an entire procedure. High fidelity simulators include animal models, cadavers and commercially available models. They have advantages of providing a more realistic environment for training, but also have disadvantages such as lack of easy availability, cost, ethical issues, veterinary assistance, anatomical variance from human organs (with animal models) and lack of bleeding and actual tissue compliance (for cadavers). The Virtual Reality simulator utilizes a computer-derived realistic virtual operative field with tactile feedback on laparoscopic instruments. The Augmented Reality simulator provides a more realistic procedure-specific operating environment, where events on the field are enhanced and supplemented [12■■,13,14].

Simulators enable residents and novice robotic surgeons to practice their skills in a nonclinical environment, any number of times, without risking the actual patients. Moreover, they provide trainees a platform to assess their performance and keep track of progress over time. Additionally, they provide an opportunity to a surgeon to refamiliarize himself with the surgical console immediately before a case as a 'warm-up' before surgery [4–10].

The simulator training can be further classified into two types – skills training and procedure-based training [4–6]. Most of the virtual reality simulators provide skills training including cutting, depth perception, hand–eye coordination, suturing and retraction. Recently, procedure-based training simulators have been reported, which can act as a bridge between formal and informal training [13,14,15■■].

In this systematic review, we have reviewed all publications in PubMed in the last 12 months using keywords: simulation, robotic training, virtual reality, augmented reality. We will discuss the current status of all existing simulators in robotic training including their advantages, disadvantages, all recently published modifications in simulators technology, assessing their place in current robotic training curriculum, along with the recent developments in simulator technology and future challenges in the simulator training for acquisition of robotic skills.

VALIDATION OF SIMULATORS

Although simulators have shown their utility over other educational tools like didactic teaching and dry lab training, they need to be validated before their effective integration into teaching and training curriculum [4–6]. Validation can be subjective and objective. The subjective validation includes face and content validity. Face validity is defined as the informal assessment of realism and feel by no experts. Content validity is defined as the formal assessment of appropriateness as a teaching tool by experts. The objective validation, which is a much more daunting task, includes construct, concurrent and predictive validity. Construct validity is defined as the ability of a simulator to discriminate experts from novices. The term 'novice' includes subjects with no experience at all in performing the procedure under study. The term 'expert' includes subjects with adequate experience in performing the procedure under study. Concurrent validity is defined as the ability to compare performance on a simulator with gold standard tests known to measure the same domain, such as a tissue or animal lab. Predictive validity is defined as the ability to predict future performance based on performance on the simulator [4–10].

VIRTUAL REALITY SIMULATORS

We found five different types of virtual reality simulators published so far in the literature.

SimSurgery Educational Platform Robot

The SimSurgery Educational Platform (SEP) Robot (SimSurgery, Oslo, Norway) is a modification of the SEP Basic laparoscopic virtual reality simulator. It replaces the simulated laparoscopic instruments with the wristed instruments found in the da Vinci robot, providing seven degrees of freedom. It does not provide three-dimensional images, fourth arm integration or performance feedback. It also does not include the following tasks: camera and clutching; needle control and driving; energy and dissection [9,16]. The experience with this simulator is not as robust as with other simulators, though it is an extremely cost-effective alternative. However, the face, content and construct validity have been proven in literature [9,16].

Robotic Surgical Simulator

The Robotic Surgical Simulator (RoSS) is another type of virtual reality simulator offering 16 modules with progressive difficulty from pinching, camera and clutch operation to tissue cutting and cautery. It is a stand-alone system mimicking da Vinci Surgical System. It helps in developing motor and cognitive skills for performing robotic surgery by providing in-vivo virtual operative steps with three levels of complexity in the form of modules for orientation, motor skills, basic surgical skills and intermediate surgical skills [17]. The face and content validity have been published for this simulator, but there is currently no literature on construct validity [4,9]. The educational impact of this simulator has been published as those trained on RoSS took less time to complete robotic dry tasks [18]

ProMIS

The ProMIS hybrid simulator (Canadian Aviation Electronics Healthcare, Canada) has a computer and a laparoscopic interface made with a plastic mannequin with a black Neoprene cover. There are three camera tracking systems to detect any instrument inside the simulator from three angles, thus recording the three-dimensional position of tips of instruments 30 times/second. It can be used for various tasks like intracorporeal suturing, precision cutting, cannulation and peg transfer, analyzing three objective parameters of time, path and smoothness [19]. The face, content and construct validity have been reported in published literature [9,19].

Mimic dV-Trainer

dV-Trainer (dVT) is a table top-sized compact system with dual-platform capability simulating both da Vinci S, Si and Xi robots. It utilizes precise modelling of robot kinematics, foot pedals and master grips. This provides trainees with a realistic representation of the da Vinci system. This provides both basic (Endowrist manipulation, camera, clutching, and troubleshooting) and advanced skills training (needle control and driving, suture and knot tying, energy and dissection) [4,7]. The face, content, construct validity and educational impact have been proven in recent published series [6,18,20–22]. Schreuder *et al.* evaluated 42 participants in three groups according to their robotic experience. Experts performed better in terms of ‘time to complete’ and ‘economy of motion’ in comparison to novices [20].

da Vinci Skills Simulator

This simulator, produced by Intuitive Surgical, can be integrated with existing da Vinci Xi or Si surgeon consoles, thus providing a practice platform to be used inside or outside the operating room, with no requirement of additional system components. This was developed in collaboration with Mimic Technologies and Simbionix and provides training modules from basic to advanced skills including Endowrist manipulation, camera and clutching, fourth arm integration, needle control and driving, energy and dissection [4,23]. The face, content and construct validity have been proven in the recent series [11,18,24–28]. Tergas *et al.* showed that training on da Vinci Skills Simulator (dVSS) resulted in significant improvement in ‘time to completion’ and ‘economy of motion’ for novices [24]. They found that autonomy of use, computerized performance feedback and ease of setup were unique advantages to dVSS, thus providing more efficient and sophisticated training in comparison to conventional dry laboratory training.

AUGMENTED REALITY SIMULATORS

These simulators provide a more realistic operating field to trainees, utilizing enhanced and supplemented events [29].

Hands-on-Surgical Training

This simulator is a mode embedded within the RoSS simulator and provides training in actual surgical cases such as radical prostatectomy, radical cystectomy, radical hysterectomy and extended lymph node dissection. It includes integrated user interaction, narrative instructions and guided movements. Hands-on-Surgical Training (HoST) was created by augmenting a real surgical procedure within a virtual reality framework utilizing audio-visual explanations and anatomically relevant illustrations of the critical steps of the procedure. The RoSS manipulators navigate the trainee through haptic-enabled cues during the procedure [13]. Chowriappa *et al.* [12▪▪] evaluated the role of augmented reality-based skills training for robot-assisted urethrovesical anastomosis in a randomized controlled trial, using HoST a technology group and a control group. They found that for 70% of participants, HoST the training experience was similar to a real surgical procedure and 75% of trainees responded that this training could improve confidence in performing a real procedure. They concluded that training with HoST in urethrovesical anastomosis improves technical skills acquisition with minimal cognitive demand.

Maestro AR

This was introduced by Mimic Technology, providing virtual instruments for interaction with anatomy in a 3D video environment. This has been designed for training novices in decision-making skills and procedure-specific skills, within the dVT simulator. The participants use virtual robotic instruments in anatomical regions collected from 3D surgical video. This simulator plans to provide training in four modules: partial nephrectomy (released May 2014), hysterectomy, prostatectomy and general surgery (to be released) by helping to identify anatomy, anticipate tissue retractions and predict regions for dissection [14]. There are no studies documenting face, content, construct, concurrent and predictive validity of this simulator, owing to its recent introduction.

RECENT DEVELOPMENTS IN CONCEPTS

Recently, more simulation models have been launched emphasizing the concept of teamwork and procedure-specific training in robotics.

Xperience Team Trainer

This simulator, available as an optional hardware complement for the dV–Trainer simulator, has been introduced to emphasize the importance of teamwork and proper coordination between console surgeon and assistant during robotic surgery. This simulator provides training simultaneously to both surgeon and bedside assistant. Thus, the bedside assistant performs basic skills exercises, promoting his psychomotor skills and rehearsal of interaction with console surgeon. It also exposes them to real-life situations in the operating room, promoting patient safety. Moreover, this team training helps in development of communication protocol in the real operating room using a well tolerated simulation environment. Moreover, it also provides proficiency-based scoring for the team and each individual [30]. However, studies regarding its face, content, construct, concurrent and predictive validity are still pending because of its recent introduction.

Tube 3 module with dV-Trainer

This simulator training emphasizes procedure-specific training, utilizing the Tube 3 module in the dVT. It helps in increasing vesicourethral anastomosis (VUA) performance, one of the most complex steps in robot-assisted radical prostatectomy. Kang *et al.* [15▪▪] recently published their experience with this module. They found that experts performed better in task time, total score, total economy of motion and number of instrument collisions in comparison with novices. Moreover, 80% of experts found this module a useful training tool to perform VUA. Thus, they reported face, content and construct validity of the Tube 3 module for practicing VUA.

RobotiX Mentor

This simulator has been introduced recently providing a realistic representation of the work space, master controllers, pedals and surgeon console of da Vinci Surgical System. It provides a 3D high-definition stereoscopic view for basic skills (robotic suturing, stapler, Fundamentals of Robotic Surgery modules) and multidisciplinary complete virtual reality procedures (vaginal cuff closure, hysterectomy modules), augmented with step-by-step video guidance and realistic representation of emergency situations and complications. The trainees are provided with performance reports with learning curve graphs utilizing simulator curricula management system [31]. However, face, content, construct, concurrent, and predictive validity of this simulator have not been proved in literature because of its recent introduction.

Table 1 shows comparison between the available simulators.

CURRENT CHALLENGES AND FUTURE PERSPECTIVES

The definitions of face, content, construct, concurrent and predictive validity need to be standardized for all simulators and future studies. Very few randomized controlled trials (RCTs) have been

Table 1. Comparison of different available simulators

	Face validity	Content validity	Construct validity	Concurrent validity	Predictive validity	Learning impact	Cross-modality correlation
SEP	Yes	Yes	Yes	No	No	No	No
RoSS	Yes	Yes	No	No	No	Yes	No
ProMIS	Yes	Yes	Yes	No	No	Yes	No
dVT	Yes	Yes	Yes	No	No	Yes	Yes
dVSS	Yes	Yes	Yes	No	No	Yes	Yes
HoST	Yes	Yes	Yes	No	No	Yes	No
Maestro AR	No	No	No	No	No	No	No
Tube-3 module	Yes	Yes	Yes	No	No	Yes	No
Xperience team trainer	No	No	No	No	No	No	No

dVSS, da Vinci Skills Simulator; dVT, dV-Trainer; HoST, Hands-on-Surgical Training; RoSS, Robotic Surgical Simulator; SEP, SimSurgery Educational Platform.

reported comparing different robotic simulators [32]. The superiority of one simulator over another has not been established so far because of a lack of these RCTs. There are no studies documenting the actual benefits of simulator training carried over to real-case performance with a surgical robot. The cost of these simulators is a significant matter of concern [4,7–9]. However, with increasing use of robotic technology and increasing competition among training devices, the future cost of these devices should come down to an affordable range. There is a need to provide more procedure-specific

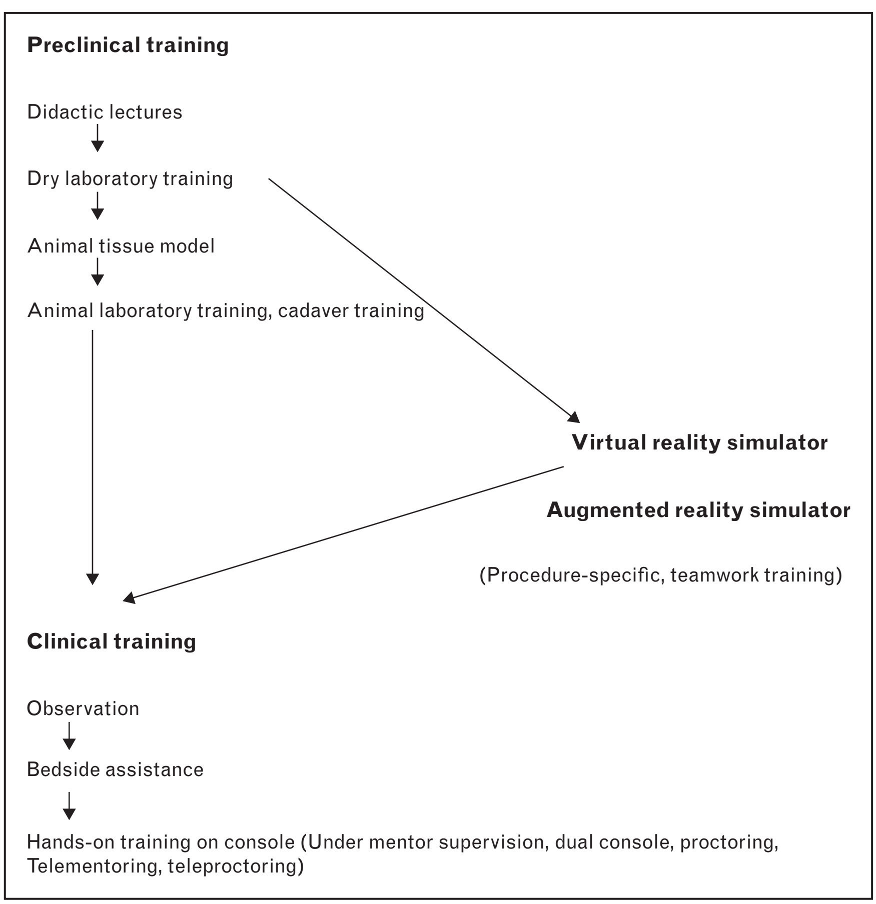


FIGURE 1. Potential role of simulators in robotics training.

training along with skills-based training in a more realistic augmented reality environment like HoST and Maestro [13,14]. Moreover, the concepts of teamworking, decision-making and communication skills should be incorporated more in simulator training by providing team-based robotic simulation environments like Xperience Team Trainer [4,7–9,30]. However, their validations have to be proved in future large prospective RCTs. Finally, there is a need for standardization for training and credentialing in robotic surgery as has been done with Fundamentals of Laparoscopy Surgery for laparoscopy in general surgery [4,7,8]. A similar standard and validated tool including simulator training and other training tools needs to be incorporated in various robotic residency and fellowship teaching curriculum (Fig. 1).

There are a few limitations of this article. First, we may have missed a few articles related to the current topic. Second, we could not discuss certain issues like cost-effectiveness, concurrent and predictive validity (tools to assess the actual benefits of simulator training carried over during real-time robotic surgery), as these issues have not been reported in published series.

CONCLUSION

The simulator training can form an integral part of credentialing and training robotic surgery of future robotic surgeons. It has the potential to decrease the learning curve for the acquisition of robotic skills. It can supplement the hands-on training clinical phase and can act as a bridge between preclinical training (didactic lectures, dry lab training, animal models) and actual hands-on clinical training without jeopardising the safety of patients. There is a need for more procedure-specific augmented reality simulator training in a cost-effective manner, utilizing advancements in computing and graphical capabilities for new innovations in simulator technology, with emphasis on both technical skills training and teamwork training. However, more RCTs involving larger numbers of participants are required to establish its cost-benefit ratio along with concurrent and predictive validity.

Acknowledgements

None.

Financial support and sponsorship

None.

Presentation at a meeting: None.

Conflicts of interest

None.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:
- ▪ of special interest
- ▪▪ of outstanding interest

1. Skarecky DW. Robotic-assisted radical prostatectomy after the first decade: surgical evolution or new paradigm. ISRN Urol 2013; 2013:157379.
2. Huang KH, Carter SC, Shih YC, Hu JC. Robotic and standard open radical prostatectomy: oncological and quality-of-life outcomes. J Comp Eff Res 2013; 2 (3):293–299.
3. Bian SX, Hoffman KE. Management of prostate cancer in elderly men. Semin Radiat Oncol 2013; 23 (3):198–205.
4. Sun AJ, Aron M, Hung AJ. Novel training methods for robotic surgery. Indian J Urol 2014; 30 (3):333–338.
5. Khan R, Aydin A, Khan MS, *et al.* Simulation-based training for prostate surgery. BJU Int 2014; doi: 10.1111/bju.12721. [Epub ahead of print]
6. Patel A, Patel M, Lytle N, *et al.* Can we become better robot surgeons through simulator practice? Surg Endosc 2014; 28 (3):847–853.
7. Buchs NC (1), Pugin F, Volonté F, Morel P. Learning tools and simulation in robotic surgery: state of the art. World J Surg 2013; 37 (12):2812–2819.
8. Liss MA (1), McDougall EM. Robotic surgical simulation. Cancer J 2013; 19 (2):124–129.
9. Abboudi H (1), Khan MS, Aboumarzouk O, *et al.* Current status of validation for robotic surgery simulators - a systematic review. BJU Int 2013; 111 (2):194–205.
10. Santomauro M, Reina GA, Stroup SP, L'Esperance JO. Telementoring in robotic surgery. Curr Opin Urol 2013; 23:141–145.
11. Rajanbabu A, Drudi L, Lau S, *et al.* Virtual reality surgical simulators-a prerequisite for robotic surgery. Indian J Surg Oncol 2014; 5 (2):125–127.
12. ▪▪ Chowriappa A, Raza SJ, Fazili A, *et al.* Augmented –reality based skills training for robot assisted urethrovesical anastomosis: a multiinstituitional randomised controlled trial. BJU Int 2014; doi: 10.1111/bju.12704. [Epub ahead of print]

This is the only randomized controlled trial reporting role of HoST™ in robotic training.

13. Simulated Surgical Systems.HoST What is HoST?2014. http://www.simulatedsurgicals.com/what-is-host.htm. Accessed September 2014.
14. Mimic Simulation Technology:Maestro AR™ http://www.mimicsimulation.com/products/maestro-ar/. Accessed September 2014.
15. ▪▪ Kang SG, Cho S, Kang SH, *et al.* The Tube 3 module designed for practicing vesicourethral anastomosis in a virtual reality robotic simulator: determination of face, content, and construct validity. Urology 2014; 84 (2):345–350.

This is the only study reporting role of Tube 3 module with dV-Trainer in procedure specific training.

16. Shamim Khan M, Ahmed K, Gavazzi A, *et al.* Development and implementation of centralized simulation training: evaluation of feasibility, acceptability and construct validity. BJU Int 2013; 111 (3):518–523.
17. Simulated Surgical Systems .RoSS™:Hardware and software. http://www.simulatedsurgicals.com/hw-sw.htm. Accessed September 2014.
18. Smith R, Truong M, Perez M. Comparative analysis of the functionality of simulators of the da Vinci surgical robot. Surg Endosc 2014. [Epub ahead of print]
19. Nagendran M, Gurusamy KS, Aggarwal R, *et al.* Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Syst Rev 2013; 8:CD006575. doi: 10.1002/14651858.CD006575.pub3.
20. Schreuder Henk WR, Persson Jan EU, Wolswijk Richard GH, *et al.* Validation of a novel virtual reality simulator for robotic surgery. ScientificWorldJournal 2014; 2014:507076.
21. Ruparel RK, Taylor AS, Patel J, *et al.* Assessment of virtual reality robotic simulation performance by urology resident trainees. J Surg Educ 2014; 71 (3):302–308.
22. Thiel DD, Patel VR, Larson T, *et al.* Assessment of robotic simulation by trainees in residency programs of the Southeastern Section of the American Urological Association. J Surg Educ 2013; 70 (5):571–577.
23. Intuitive Surgical.da Vinci Skills Simulator. http://www.intuitivesurgical.com/products/skills_simulator/. Accessed September 2014.
24. Tergas AI, Sheth SB, Green IC, *et al.* A pilot study of surgical training using a virtual robotic surgery simulator. JSLS 2013; 17 (2):219–226.
25. Alzahrani T, Haddad R, Alkhayal A, *et al.* Validation of the da Vinci Surgical Skill Simulator across three surgical disciplines:A pilot study. Can Urol Assoc J 2013; 7 (7–8):e520–e529.
26. Brinkman WM, Luursema JM, Kengen B, *et al.* da Vinci skills simulator for assessing learning curve and criterion-based training of robotic basic skills. Urology 2013; 81:562–566.

27. Lyons C, Goldfarb D, Jones SL, *et al.* Which skills really matter? Proving face, content, and construct validity for a commercial robotic simulator. Surg Endosc 2013; 27:2020–2030.

28. Hung AJ, Jayaratna IS, Teruya K, *et al.* Comparative assessment of three standardized robotic surgery training methods. BJU Int 2013; 112:864–871.

29. Stegemann AP, Ahmed K, Syed JR, *et al.* Fundamental skills of robotic surgery: A multiinstitutional randomized controlled trial for validation of a simulation-based curriculum. Urology 2013; 81:767–774.

30. Mimic Simulation Technology:Xperience Team Trainer^{TM} http://www.mimicsimulation.com/products/xperience/. Accessed September 2014.

31. Simbionix^{TM}: RobotiX Mentor^{TM}. http://simbionix.com/simulators/robotix-mentor/. Accessed September 2014.

32. Korets R, Mues AC, Graversen J, *et al.* Comparison of robotic surgery skill acquistion between DV-Trainer and da-Vinci surgical system: a randomized controlled study. J Urol 2011; 185:e593.

Please select Print from the file menu to print your Abstract.

43rd AAGL Global Congress on Minimally Invasive Gynecology

Abstract Number: 450754
Presenting Author: Mireille Truong, Fellow
Correspondence Contact:Mireille Truong, MD
Institution: Columbia University Medical Center
Address: 200 W 60th Street, apt 18H
City/State/Zip/Country: New York, NY, 10032, United States
Phone: 804-363-9436 Fax: E-mail: mireille.truong@gmail.com

DISCLOSURE INFORMATION

Author: Mireille Truong

I, the undersigned, declare that I (or an immediate family member) do NOT have a financial arrangement or affiliation with any corporate organization offering financial support or grant monies for the Global Congress, a continuing medical education program, or with any corporate organization with a financial interest in the abstract that I am submitting.

Off Label: I do not have any "off label" uses in my presentation

Signature: I verify, that all of the above disclosures are true.

Signature Name: Mireille Truong

Signature Date: May 1, 2014 2:06 am

Author: Alyssa Tanaka*

I, the undersigned, declare that I (or an immediate family member) do NOT have a financial arrangement or affiliation with any corporate organization offering financial support or grant monies for the Global Congress, a continuing medical education program, or with any corporate organization with a financial interest in the abstract that I am submitting.

Off Label: I do not have any "off label" uses in my presentation

Signature: I verify, that all of the above disclosures are true.

Signature Name: Alyssa Tanaka

Signature Date: May 1, 2014 2:07 am

Author: Khara Simpson*

I, the undersigned, declare that I (or an immediate family member) do NOT have a financial arrangement or affiliation with any corporate organization offering financial support or grant monies for the Global Congress, a continuing medical education program, or with any corporate organization with a financial interest in the abstract that I am submitting.

Off Label: I do not have any "off label" uses in my presentation

Signature: I verify, that all of the above disclosures are true.

Signature Name: Khara Simpson

Signature Date: May 1, 2014 2:09 am

Author: Arnold Advincula*

I, the undersigned, declare that I (or an immediate family member) do NOT have a financial arrangement or affiliation with any corporate organization offering financial support or grant monies for the Global Congress, a continuing medical education program, or with any corporate organization with a financial interest in the abstract that I am submitting.

Off Label: I do not have any "off label" uses in my presentation

Signature: I verify, that all of the above disclosures are true.

Signature Name: Arnold Advincula

Signature Date: May 1, 2014 2:31 am

Author: Roger Smith*

I, the undersigned, declare that I (or an immediate family member) do NOT have a financial arrangement or affiliation with any corporate organization offering financial support or grant monies for the Global Congress, a continuing medical education program, or with any corporate organization with a financial interest in the abstract that I am submitting.

Off Label: I do not have any "off label" uses in my presentation

Signature: I verify, that all of the above disclosures are true.

Signature Name: Roger Smith

Signature Date: May 1, 2014 2:08 am

QUESTIONNAIRE INFORMATION

1. Application for: Abstract or Poster
2. Type of Abstract: Educational
3. Category: 2. Education
4. Type of Award: Jerome J. Hoffman Award for Best Postgraduate Paper on MIGS (Residents and Fellows in training and from any program may apply)
5. Is this abstract original? Yes

Title: A prospective randomized controlled comparative study on surgical training methods and impact on surgical performance: virtual reality robotic simulation vs didactic lectures

Mireille Truong, MD, mireille.truong@gmail.com 8043639436[1], Alyssa Tanaka, MS, alyssa.tanaka@flhosp.org[2], Khara Simpson, MD, kmsimpmd@yahoo.com[2], Arnold Advincula, MD, aa3530@cumc.columbia.edu[1] and Roger Smith, PhD, roger.smith@flhosp.org[2]. [1]Obstetrics and Gynecology, Columbia University Medical Center, New York, NY, United States, 10032 and [2]Research Department, Florida Hospital Nicholson Center at Celebration Health, Celebration, FL, United States, 34747.

Objective: To compare two learning methods, structured didactic lecture versus virtual reality surgical simulation, and evaluate their effects on surgical performance.

Design: Prospective randomized trial

Setting: Community teaching hospital and robotic gynecologic surgery courses

Patients: Medical students (n=6), residents (n=14), fellows (n=28), and attendings (n=76) were enrolled

Interventions: Subjects, based on surgical experience level, were randomly assigned to either the Simulation Group (SG), which rehearsed on the dv-Trainer robotic simulator; or to the Didactic Group (DG), which received a structured lecture. After completing a written post-preparation test, all participants performed a video-recorded cystotomy repair on a live porcine model. Time and performances were compared using GEARS and independently

developed error and task metrics.

Measurements & Main Results: Total n= 125 (52 novices, 42 intermediates, 27 experts). Both groups (DG n=64; SG n=61) were similar in age, gender, role, and total number of robotic cases. The majority of subjects had prior robotic simulator use, DG>SG (p = 0.04). Mean cystotomy repair time was similar in both groups (DG=224min; SG=219min, p=0.83). The overall performances between SG and DG were not significantly different (p value= 0.18 – 0.83) but when controlled for experience level, SG (vs DG) novices had more errors and lower task assessment scores (p value 0.03 and 0.05). No differences were noted between learning groups amongst intermediates and experts.

Conclusions: Simulation and didactic approaches both offer certain advantages for surgical training. Although differences were not seen between the two training methods, many factors may have influenced these results such as simulator training type, task complexity and length; realism of simulation exercise to actual procedure; and differences in information presented with each method. Both cognitive and psychomotor skills are required for surgical competence; therefore the effect of the combination of both modalities rather than a single modality for surgical training should be further explored.

Close Window

Crowd-Sourced Assessment of Technical Skills (C-SATS): Differentiating Animate Surgical Skill Through the Wisdom of Crowds

HOLST, Daniel*; KOWALEWSKI, Timothy M., PhD††; WHITE, Lee W., PhD**; BRAND, Timothy C., MD***; HARPER, Jonathan D., MD†; SORENSON, Mathew D., MD†; TRUONG, Mireille†††, MD, SIMPSON, Khara†††, MD, TANAKA, Alyssa, †††, SMITH, Roger, PhD.†††; LENDVAY, Thomas S., MD†

*University of Washington School of Medicine; **Stanford University School of Medicine; ***Madigan Army Medical Center, Department of Urology; †University of Washington, Department of Urology; ††University of Minnesota, Department of Mechanical Engineering; †††Florida Hospital Nicholson Center.

ABSTRACT:

Background: Objective quantification of surgical skill is imperative as we enter a healthcare environment of quality improvement and performance-based reimbursement. The gold standard tools are infrequently used due to time-intensiveness, cost-inefficiency, and lack of standard practices. We hypothesized that valid performance scores of surgical skill can be obtained through crowd-sourcing.

Methods: Twelve surgeons of varying robotic surgical experience performed live porcine robotic-assisted urinary bladder closures. Blinded video-recorded performances were scored by expert surgeon graders and by Amazon.com Mechanical Turk crowd-sourcing crowd workers using the Global Evaluative Assessment of Robotic Surgery (GEARS) tool assessing five technical skills domains. Seven expert graders and 50 unique Mechanical Turkers (each paid $0.75/survey) evaluated each video. Global assessment scores were analyzed for correlation and agreement.

Results: Six-hundred Mechanical Turkers completed the surveys in under 5 hours, while 7 surgeon graders took 14 days. The duration of video clips ranged from 2-11 minutes. The correlation coefficient between the Turkers and expert graders' scores was 0.95 and Cronbach's alpha was 0.93. Inter-rater reliability among the surgeon graders was 0.89.

Conclusion: Crowd-sourcing surgical skills assessment yielded rapid, inexpensive agreement with global performance scores given by expert surgeon graders. The crowd sourcing method may provide surgical educators and medical institutions with a boundless number of procedural skills assessors to efficiently quantify technical skills for use in trainee advancement and hospital quality improvement.

INTRODUCTION:

The healthcare environment is shifting towards performance-based reimbursement and focusing on quality improvement. A 2000 study from the Agency for Healthcare Research and Quality showed that surgical mortality is among the top 10 causes of death in the United States.[REF] While not all deaths from surgery were due to technical error in this particular report, another study evaluating the role of surgical trainee's in malpractice claims data showed that errors in manual technique were present in 56% of all cases.[REF]

Recent literature has shown that blinded video assessments of technical performances among experienced laparoscopic surgeons directly correlates with patient outcomes.[REF] Subsequently, efforts have been made to adopt methods for evaluating technical skill using global surgical performance-rating scales, like the Objective Structured Assessment of Technical Skills (OSATS) tool, and its derivatives like GEARS or GOALS tools.[REF] These methods are validated and are considered the gold-standard for evaluating surgical technical skill, but they are resource and time intensive, require the time of busy surgeons to evaluate videos, and are thus infrequently used. A cheaper, faster, less-biased means of assessing technical skill is needed.

Crowd-sourcing is a means of accomplishing tasks through the work of decentralized, independent groups of people who are generally non-experts in the task that is being accomplished.[REF] The advent of the internet has enabled a global labor market ready to do various tasks/surveys to help solve problems, from helping blind mobile phone users navigate their surroundings,[REF] to discovering complex protein folding structures.[REF] In recent studies, crowds have been shown to be as effective as expert surgeons at evaluating surgical technical skill in a dry lab setting.[Chen et al. and Holst et al.] Not only did the crowds perform as effectively as the expert surgeons in providing skill assessment, but the costs, efficiency, and practicality of use were all improved with crowd graders vs. expert surgeon graders. The major limitation of these studies was that the surgical tasks being assessed were dry-lab tasks. Thus, no real tissue was being manipulated in the study leaving questions regarding whether non-experts can appreciate the subtlety of real surgery. In this study, we hypothesize that crowd-sourcing can be used to obtain valid performance grading of surgical technical skill on real, living, viable tissue.

MATERIALS AND METHODS

After institutional review board approval, two reviewer groups were recruited for this study: Amazon.com Mechanical Turk™(Amazon.com, Seattle, WA) users and expert faculty surgeon graders, who have expertise in robotic surgery. Surgeons were recruited via email and crowd-workers were recruited through the Amazon.com Mechanical Turk™ platform. Six-hundred pre-qualified Mechanical Turkers™ were recruited for the study (Figure 1). In order to qualify for the study, the crowd-workers had to have previously completed 100 or more Human Intelligence Tasks (HITs), the task unit used by Mechanical Turk™, and must have had a greater than 95% approval rating as qualified by the Mechanical Turk™ as described in Chen et al.[REF] These workers were identified only by a unique, anonymous user identification code provided by Mechanical Turk™ and no other information was known about them (gender, age, sex, ethnicity, etc.). Each crowd-worker was compensated $0.75 USD for participating. The expert faculty surgeon grader group consisted of seven experienced robotic surgeons, who have all practiced as attending surgeons for a minimum of three years with predominantly minimally invasive surgery practices and who were familiar with evaluating surgical performances by video analysis. The expert surgeons did not receive monetary compensation. All graders were required to be over 18 years old.

A surgical skill assessment survey was developed and hosted online on a secure server. The survey consisted of an initial qualification question in which the crowd reviewers were shown two videos of a pair of surgeons performing a Robotic Fundamentals of Laparoscopic Surgery (RFLS) block transfer task displayed side-by-side (Figure 2). The surgeon in the left video performed the task with a high level of skill while the surgeon in the right video performed the task with intermediate level of skill. The skill level was based on published benchmark metrics for this particular task.[REF] Crowd-workers were asked to pick the video with the higher level of skill. Crowd-workers who incorrectly answered the qualification question were excluded from the analysis but were still remunerated. Additionally, the survey also contained an attention question to ensure the assessors were actively paying attention and those crowd-workers who incorrectly answered were also excluded from the analysis. (FIGURE 1)

For the second part of the survey, we obtained twelve recorded videos of twelve different surgeons of varying skill level performing live porcine robotic-assisted urinary bladder closures. (Figure 3). No identifying information of the surgeons performing the bladder closures was present and all graders (crowd and expert) were blinded to the identity of the surgeons. The length of the videos ranged between two minutes and eleven minutes and the average length was four and a half minutes. The videos were uploaded to the online survey which incorporated five domains from the Global Evaluative Assessment of Robotic Skills (GEARS) validated robotic surgery rating tool – bimanual dexterity, depth perception, efficiency, force sensitivity, robotic control.[REF] (FIGURE 4) Fifty unique Mechanical Turk™ crowd-workers and seven expert surgeons graded each video based on the five skills domains. We chose fifty crowd-workers per video based on a previous internal analysis of data collected by our team (Chen et al 2014) which found 30-50 crowd responses sufficient to achieve satisfactory agreement with expert grades.[REF].

Composite performance scores were tallied by summing the Likert ratings across the five domains with a scale of 5-25. The mean of the crowd-worker's composite scores was compared to that of the expert faculty surgeons' and assessed for correlation using Cronbach's Alpha statistic to assess the degree of concordance. (Table 1) According to common practice, levels above 0.9 indicate "excellent agreement," down to 0.7 "good agreement" and levels below 0.5 indicate "poor, unacceptable" levels of agreement.[REF].

RESULTS:

After excluding crowd-workers who failed the attention or discrimination question, we were left with valid grades from 487 Mechanical Turk™ crowd-workers. (FIGURE 1) It took 4 hours 28 minutes to receive all crowd-worker grades for the 12 videos. In comparison, it took 14 days to receive grades from all 7 expert surgeons. Composite scores given by both the crowds and experts are shown in Table 1. Inter-rater reliability between the surgeons and crowd was 0.93 using Cronbach's Alpha statistic, which indicates "excellent" agreement. (Table 1) The linear relationship between the surgeon grades and crowd grades is shown in Figure 3. The R^2 value is 0.91.

DISCUSSION:

The current gold standard - OSATS-like methods for objectively assessing surgical skill - continues to be underutilized due to cost, resource-intensiveness, and the lag-time for return of results. [REF] Feedback is most effective if given immediately or near real-time so existing OSATS practices tend to be deficient outside of an academic research project. [REF DARZI PAPER] Because expert assessment may have significant variability in the absence of an 'agreement workshop' and because mentor bandwidth precludes frequent iterative trainee objective technical skills assessment, alternative methods to assist in these goals are required. Additionally, OSATS may not be that 'objective' as the reviews tend to be done by reviewers who are within the same institution as the reviewees.[REF].

In Holst et al. and Chen et al., it was noted that C-SATS was not designed to replace one-on-one instruction and evaluation in the setting of residency training, but may provide an adjunct. Traditional methods of instruction and feedback are invaluable because they offer content expertise and transfer information about the nuances of surgery that could not be yielded by crowds.[REF] C-SATS, however, may have a role in rapidly triaging trainees with deficiencies and then allowing mentors to target valuable training resources to these deficiencies as opposed to teaching all trainees with the same curricula. Feedback from crowds may be obtained rapidly enough to provide guidance between surgical cases or between days in the operating room.

C-SATS has been used in a residency training environment which is ideally suited to this method because of the controlled, learner-centered nature of residency. In Holst et al., they showed that crowds can identify differences in Urology resident training levels and that crowd-sourcing is a practical, effective way of providing feedback to Urology residents in near real-time. [REF Holst J Endourology paper] The major limitation of that study, however, was that all tasks evaluated were dry-lab tasks. In a setting of resident work-hour restrictions, surgical trainees are spending more time in simulation labs to refine their technical skills, and thus it is important that crowds can evaluate these dry lab tasks quickly. However it is vital to prove that crowds can judge technical skill being performed on real, live tissue as opposed to dry-lab materials. Animate surgery better approximates real human surgery, thus our hypothesis needed to be tested in this environment as a 'next step' in validating C-SATS. With no knowledge of relevant anatomy, crowds provided extremely rapid and accurate feedback in comparison to expert graders.

A limitation of this study is that only one type of live-tissue performance was assessed and the surgery was still a very controlled environment being a porcine lab. In addition, all videos assessed were relatively short (averaging under 5 minutes in length). It remains to be seen if crowd evaluators can continue to provide effective grading across a range of live-tissue surgeries with varying lengths. Future studies aim to include videos across a range of surgical approaches such as laparoscopic and open surgeries. Additional validation is needed before C-SATS is imbedded into training centers, and evidence that crowds can evaluate live-tissue surgery adds to the growing body of evidence for the value of this adjunctive objective assessment tool.

Another limitation to this study is that the performances assessed were from a wide range of surgical skill levels from robotic faculty to novice trainees. Thus, the skill 'effect-size' may have been disparate enough for lay people to easily see differences. It is arguable that if the cohort of performers were of more similar skill levels, it would require expert observers to discriminate the smaller technical skills differences. Resident training environments where the skills of the trainees vary significantly are ideally suited to using this methodology. Additional studies will be needed to test C-SATS on cohorts of surgeons who have similar skills.

CONCLUSIONS: We demonstrate that crowd-sourcing basic surgical skills of animate surgery compares favorably to a panel of expert surgeon assessors and is faster than the experts - providing large-volume feedback in a matter of hours. Utilizing crowd-sourcing as a means to assess technical surgical skills provides an inexpensive, scalable, rapid and effective way to evaluate live-tissue procedures, paving the way for further validation in human surgery. Ultimately, C-SATS assessments will need to be linked to clinical outcomes to gain confidence that presumably non-medically-trained crowds of people can accurately ascribe surgical skill.

From Design to Conception: An Assessment Device for Robotic Surgeons

Alyssa Tanaka, M.S.
Florida Hospital Nicholson Center
Celebration, FL
Alyssa.tanaka@flhosp.org

Manuela Perez, M.D.
University Hospital of Nancy
Nancy, FR
m.perez@chu-nancy.fr

Mireille Truong M.D., Khara Simpson M.D.
Columbia University Medical School
New York, NY
Mireille.truong@gmail.com, kmsimpmd@yahoo.com

Gareth Hearn, Roger Smith, Ph.D.
Florida Hospital ISA, Florida Hospital NC
Orlando, FL, Celebration, FL
Gareth.hearn@flhosp.org, roger.smith@flhosp.org

ABSTRACT

The daVinci Surgical System offers surgeons improved capabilities for performing complex minimally invasive procedures; however, there is no standardized assessment of robotic surgeons and a need exists to ensure that a minimal standard of care is provided to all patients. The Department of Defense and governing surgical societies convened consensus conferences to develop a national initiative, resulting in a curriculum called the Fundamentals of Robotic Surgery (FRS). FRS is comprised of an online curriculum and a psychomotor skills dome.

This paper describes the production process used to create a psychomotor skills assessment device - the FRS Dome. The device was designed to measure the essential skills that are required of any robotic surgeon and to provide a basis upon which to grant or deny privileging with the robot. It was constructed to test seven tasks of manual dexterity: Docking, Ring Tower Transfer, Knot Tying, Suturing, 4th Arm Cutting, Puzzle Piece Dissection, and Energy Dissection.

The initial design of the device was created by a committee of experienced minimally invasive surgeons, with a background in testing protocols and materials. The design was rendered in computer animation, which kick-started a prototyping effort with physical materials. These included platinum cure silicone approximating human tissue and a 3D polyjet printer for the structural framework. Usability testing was conducted and iterative modifications were made to improve ergonomics, standardization, and cost requirements. Final CAD diagrams and specifications were created and distributed to medical and simulation companies for both physical and digital manufacturing. This development process demonstrates the evolution of a simulation and a physical testing device based on international expert consensus. The specifications are open source, allowing competitive production and future iterations. The goal of this paper is to discuss how this device evolved from an idea to a manufactured product and a digital simulation.

ABOUT THE AUTHORS

Alyssa D.S. Tanaka, M.S. is a Systems Engineer at Florida Hospital's Nicholson Center. Her research work focuses on robotic surgery simulation and effective surgeon training. Her current projects include rapid prototyping of surgical education devices, the validation of a robotic surgical curriculum and evaluation of robotic simulation systems. She is a Modeling and Simulation PhD student at the University of Central Florida and previously earned a M.S. in Modeling and Simulation, Graduate Simulation Certificate in Instructional Design, and a B.S. in Psychology and Cognitive Sciences from the University of Central Florida.

Manuela Perez, M.D. is a practicing General Surgeon at the University Hospital of Nancy-France, where she also serves as an Assistant Professor in General Surgery and Anatomy. Dr. Perez has been practicing medicine for 14 years and graduated with her PhD in Robotic Surgery, with a thesis entitled "Telesurgery: From Training to Implementation." Currently, she is working as a Research Fellow at the Florida Hospital Nicholson Center and working under a grant from the Department of Defense researching various aspects of Telesurgery.

Mireille Truong, M.D. is a Minimally Invasive Gynecology Fellow at Columbia University, where she serves as an assistant attending and clinical instructor of Obstetrics and Gynecology (OB/GYN). During her residency in OB/GYN at the University of Illinois at Chicago, she received a number of awards and honors, including Administrative Chief Resident, the American Association of Gynecologic Laparoscopists' Special Resident in Minimally Invasive Gynecology Award and Best Overall Excellence in Gynecologic Care Award. She has dedicated her time to education of medical students and physicians via an array of academic and teaching appointments, professional organization board positions, peer-reviewed publications and presentations at various national and international conferences. Her research interests include surgical education, robotic simulation and minimally invasive gynecologic surgery.

Khara Simpson, M.D. is a second year fellow in minimally invasive surgery at Columbia University, where she serves as an assistant attending and instructor of obstetrics and gynecology. She completed her medical school education at Howard University College of Medicine where she was inducted into the Alpha Omega Alpha honor medical society. Following, she completed her OB/GYN residency at Johns Hopkins University and served as administrative chief resident. She recently completed a one year research fellowship at the Florida Hospital Nicholson Center focusing on robotic surgery simulation. Her additional research interests include resident education and simulation training, and best practices to promote cost effective care.

Gareth Hearn a Mechanical Engineer at the Institute for Surgical Advancement, part of Florida Hospital Orlando. He is responsible for the Prototype Design Lab which is focused on minimally invasive surgical devices and positioned to accelerate the product development cycle. He has 8 years' experience working in the Dept. of Defense training and simulation industry. He is pursuing a Master's in Systems Engineering at the University of Florida and has previously earned a B.S. in Mechanical Engineering from UF.

Roger Smith, Ph.D. is an expert in the development of simulation devices and training programs. He has spent 25 years creating leading edge simulators for the Department of Defense and Intelligence agencies, as well as accredited methods for training with these devices. He is currently the Chief Technology Officer for the Florida Hospital Nicholson Center where he is responsible for establishing the technology strategy and leading technology implementation. He has served as the CTO for the U.S. Army PEO for Simulation, Training and Instrumentation (PEO-STRI); VP and CTO for training systems at Titan Corp; and Vice President of Technology at BTG Inc. He holds a Ph.D. in Computer Science, a Doctorate in Management, and an M.S. in Statistics. He has published 3 professional textbooks on simulation, 10 book chapters, and over 100 journal and conference papers. His most recent book is *Innovation for Innovators: Leadership in a Changing World*. He has served on the editorial boards of the *Transactions on Modeling and Computer Simulation* and the *Research Technology Management* journals.

From Design to Conception: An Assessment Device for Robotic Surgeons

Alyssa Tanaka, M.S.
Florida Hospital Nicholson Center
Celebration, FL
Alyssa.tanaka@flhosp.org

Manuela Perez, M.D.
University Hospital of Nancy
Nancy, FR
m.perez@chu-nancy.fr

Mireille Truong M.D., Khara Simpson M.D.
Columbia University Medical School
New York, NY
Mireille.truong@gmail.com, kmsimpmd@yahoo.com

Gareth Hearn, Roger Smith, Ph.D.
Florida Hospital ISA, Florida Hospital NC
Orlando, FL, Celebration, FL
Gareth.hearn@flhosp.org, roger.smith@flhosp.org

INTRODUCTION AND BACKGROUND

Robotic surgery has been established as an innovative approach in surgery due to a telemanipulator device, which introduced a new dimension into surgical tools. This device allows surgeons to manipulate robotic arms from a remote console to perform complex surgical procedures. Robotic surgical systems overcome laparoscopic limitations and facilitate the performance of minimally invasive surgery due to 3D vision, 7-degree-of-freedom instruments, tremor abolition, motion amplification, and stabilization of the camera (Patel et al., 2013; Hubens, Coveliers, Balliu, Ruppert, & Vaneerdeweg, 2003; Blavier, Gaudissart, Cadière, & Nyssen, 2007). The system also offers 10x magnification, wristed instruments, and a third working arm. Currently, the only system is Intuitive's da Vinci Surgical System (Figure 1).

Figure 1. da Vinci Surgical System

Robotic surgery has demonstrated safety and effectiveness for urologic, gynecologic, ENT, and complex general surgery procedures (Barbash, Friedman, Glied, & Steiner, 2014; Serati et al., 2014; Maan, Gibbins, Al-Jabri, & D'Souza, 2012; Luca et al., 2013; Zureikat et al., 2013). Exponential growth of minimally invasive procedures, particularly robotic-assisted procedures, raises the question of how to assess robotic surgical skills. This device also introduces a specific need for training and certification to ensure a minimal standard of care for all patients. Some institutions have attempted to develop and validate robotic training in regards to specific specialties (Chitwood et al., 2001; Geller, Schuler, & Boggess, 2011; Grover, Tan, Srivastava, Leung, & Tewari, 2010; Chowriappa et al., 2014; Jarc & Curet, 2014); however, the lack of a national standard has pushed surgical societies (e.g. the Society of American Gastrointestinal and Endoscopic Surgeons and Society of Robotic Surgery) to develop a unified approach and standard for robotic skills training (Zorn et al., 2009).

To develop a comprehensive model for robotic surgery, the Department of Defense, Veterans Administration, and fourteen surgical specialty societies convened multiple consensus conferences to create the Fundamentals of Robotic Surgery (FRS) curriculum. A similar education and training initiative was implemented for use in laparoscopic surgery, which resulted in the Fundamentals of Laparoscopic Surgery (FLS). FRS Conference participants included more than 80 subject matter experts (SMEs), consisting of surgeons, psychologists, engineers, simulation experts, and medical educators (Smith, Patel, Chauhan, & Satava, 2013).

The committee's vision of FRS was driven by two main goals: to ensure a perfect understanding of the basics of robotic surgery and to develop a psychomotor skills program that focused on basic robotic tasks. The intended users for this program are novice robotic surgeons, who could be residents or fellows and attending surgeons who have never used the robotic system. The committee began by outlining outcomes measures and metrics, which touched on the essential cognitive, psychomotor, and team training skills. This resulted in a prioritized matrix of 25 robotic surgery concepts, which is the core material used in the design and development of the FRS Curriculum (Smith, Patel, Satava R, 2013). Two assessment tools were created: an online curriculum for knowledge and team training skills and a device for psychomotor skill training and evaluation (Levy, n.d.).

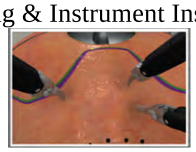
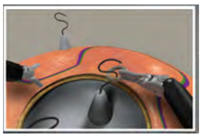
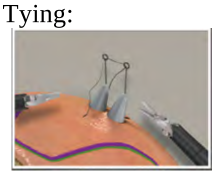
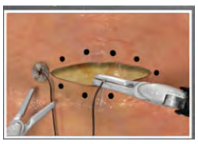
This paper discusses the process for designing and creating the physical device, known as the FRS dome. The purpose is to share the evolution of an idea to a usable device. The dome was conceived by experts who identified a clear need for robotic education and collectively developed a solution to fill the gap. The medical field is a constant progression of new concepts, devices, and technology. This paper also outlines the framework for which others can develop and introduce new concepts in medicine and other domains.

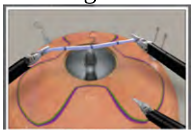
BRAINSTORMING AND CONCEPT DEVELOPMENT

Exercise Development

Of the 25 FRS concepts, 16 are directly linked with psychomotor skills. The FRS committee members then identified seven exercises that incorporated all 16 skills. These exercices include docking and instrument insertion, tower transfer, knot tying, railroad track, 4th arm cutting, puzzle piece dissection, and vessel energy dissection (Table 1). *Docking and instrument insertion* is an essential and unique robotic skill to begin a procedure. Failure at this stage of the procedure can compromise the surgery. Ring *Tower transfer* is a non-surgical exercise that introduces the utilization of endowrist manipulation and the 7 degrees of freedom to surgeons. *Knot tying* and *railroad track* are the base of a suturing exercise. The technology introduced in the wristed instruments facilitates the performance of these tasks. *4th arm cutting* is another task specific to robotics, which tests surgeon's autonomy. The 4th arm allows surgeons to manage three instruments by using a foot pedal to switch between working arms. *Puzzle piece* and *vessel energy dissection* are critical tasks, which incorporate complex articulation of instruments and application of energy (i.e. cauterization and cutting).

Table 1: Description of the basic psychomotor skills attached to the seven FRS tasks.

Exercises	Skills
Task1: Docking & Instrument Insertion: 	- Docking - Instrument insertion - Eye-hand coordination - Operative field of view
Task 2: Ring Tower Transfer: 	- Eye-hand coordination - Camera navigation - Clutching - Wrist articulation - A-traumatic handling
Task 3: Knot Tying: 	- Knot tying - Suture handling - Eye-hand coordination - Wrist articulation
Task 4: Railroad Track: 	- Needle handling & manipulation - Wrist articulation - A-traumatic handling - Eye-hand coordination

Task 5: 4th Arm Cutting: 	- Multiple arm control & switch - Cutting - A-traumatic handling - Eye-hand coordination
Task 6: Puzzle Piece Dissection: 	- Sharp and blunt dissection - Cutting - A-traumatic handling - Eye-hand coordination - Wrist articulation
Task 7: Vessel Energy Dissection: 	- Energy sources use - Sharp dissection - Cutting - Multiple arm control - A-traumatic handling - Eye-hand coordination

Device Development

The FRS committee envisioned all of the exercises contained on the outer surface of a single device. This would allow for the exercises to be administered quickly and easily, incur less cost, and ensure uncomplicated storage and transportation. The semi-spherical form (i.e. the dome), was quickly decided on as a shape which would integrate with the current robotic system. They depicted their ideas through simple drawings and crude models made from materials found on hand. During initial design planning, conference participants experimented with a variety of arrangements of the exercises on the dome.

A final sketch was developed and delivered to a 3D digital artist to create static pictures of the device, along with an animation of the performance of each exercise. The CGI provided the first formal images of the dome, which gave life to the device and proved feasibility. The realistic animations showed the exercises being performed and gave committee members a visual concept of how the device would function (Figure 2).

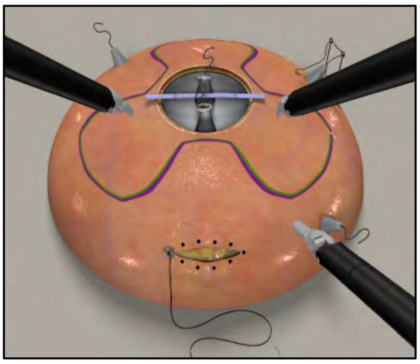


Figure 2. The initial 3D graphic FRS dome design

PROTOTYPING

The prototyping process began using the ideas developed in the design meeting and the CGI. This process would prove to be fundamental in confirming the design expectations. It was essential to determine if a single device could physically house all of the exercises effectively, if the planned architecture was compatible with the robotic system, and if the outcomes of the exercises could be measurable and reproducible.

Low-fidelity Prototypes

Low-fidelity prototypes (LFPs) were created using simple and inexpensive materials. None of the materials used in the LFPs were intended for inclusion in a final product. These materials were chosen because they were readily available, inexpensive, and easy to manipulate to test fit and function. These materials allowed rapid trial and error testing of the technical aspects, clarifying requirements, and proving usability. The testing of the LFPs was performed using the da Vinci Surgical System and was video recorded. These recordings were sent to FRS committee members to provide their feedback. Each LFP resulted in multiple improvements to the designs, which were tested on subsequent prototype versions.

The base model of the LFPs was created using half of an 8” Styrofoam sphere as the support structure, yellow felt material as the fat layer, a latex swimming cap for the skin layer, and straws for the embedded vessels. The base of the towers was constructed using synthetic foam blocks carved into a cone shape (Figure 3). The exercise patterns were drawn onto the surface using a permanent marker.

Figure 3. Base of Low Fidelity Prototypes

The LFPs evolved over six iterations, all of which introduced design improvements (Figure 4). At the earliest phase in LFP testing, it was quickly realized that the dome size was too large to fit under the robot arms appropriately. So, the dome size was decreased from 8” to 7”. Another modification made early in the LFP development was to change the 4th arm cutting band from a rigid tube to an elastic band. This allowed for the user to adequately stretch the band prior to each cut.

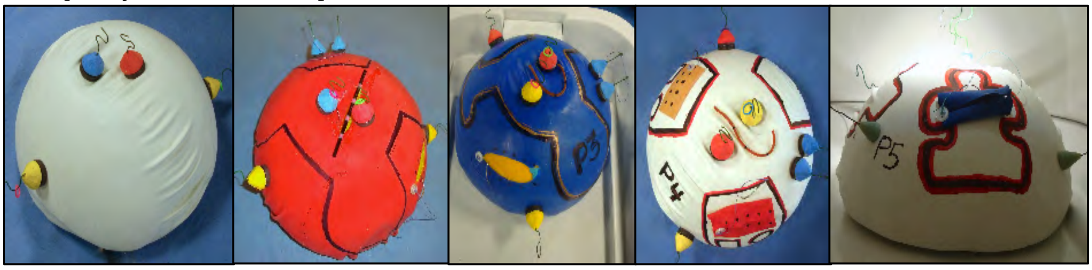


Figure 4. Iterations of LFPs

The suturing and dissection exercises involved the most modifications during the LFP stages. The original cloverleaf shape, used for the dissection exercise, was found to be too large and did not allow for the surgeons to access the section of the shape that was located on the backside of the dome. The size of the pattern was reduced; however, this did not mitigate the accessibility issue. The team experimented with other options, such as splitting the clover leaf into three sections and adding smaller shapes to the center of the cutting area. This design was not practical because once the smaller shapes were cut, the latex receded and inhibited surgeons from cutting the surrounding shape.

Eventually, the dissection shape evolved to a puzzle piece that incorporated all of the prerequisites for the dissection exercise (i.e. an accessible shape and a complex design). By using this compact pattern it became clear that all exercises could be grouped into an area covering only one third of the surface of the dome. This opened the opportunity to replicate the cluster of exercises three times on the surface, reducing the materials and costs for repeatedly practicing with the device. Another obstacle was to build the suturing exercise with the adequate materials and placements, to ensure a realistic feeling of suturing. Originally, the incision was made into the latex swim cap, however the latex would tear away and recede after the incision was cut in this model. Two versions of the suture module were experimented with: an embedded silicone and an external latex model. Eventually the embedded silicone model was chosen as the most realistic and practical for the exercise. Ultimately, the basic structural changes found in the low-fidelity prototyping were:

- The dome base needed to be reduced to 7”
- The dome base needed to be substantial in weight to keep from moving under the force of the robot
- A smaller, yet equally complex dissection shape was necessary
- The exercise sets could be grouped to allow them to be repeated on the surface of a single dome

- The magnets which held the towers to the dome needed to be of sufficient strength to hold through the layers of fat and skin

High-Fidelity Prototypes

The high-fidelity prototypes (HFPs) were made using higher quality, custom materials. These materials had the desired qualities of the final product and could be used as a basis for the large scale manufacturing process. The styrofoam base from the LFPs was replaced with a support structure that was printed using a 3D polyjet printer (Figure 5). A polyjet type 3D printer works similarly to an inkjet printer in that it distributes layers of polymer to build the desired design, which is cured by UV light. This type of printer was chosen because of the versatility allowed by printing multiple materials at once. Also, the jet lays $16\mu m$ layers of liquid polymer, which gives printed parts a finer resolution. Using this printer, a dome shell with a lid was created. The shell and lid had divots covering the surface, allowing for magnets to be moved to many different placements on the dome during design experiments. A small jig was also created using the 3D printer. Prior to the creation of the jig, the wires were made by hand, but the jig enables the standardized creation of the S-shaped and I-shaped tower wires. The price to print these items was approximately $1,000.

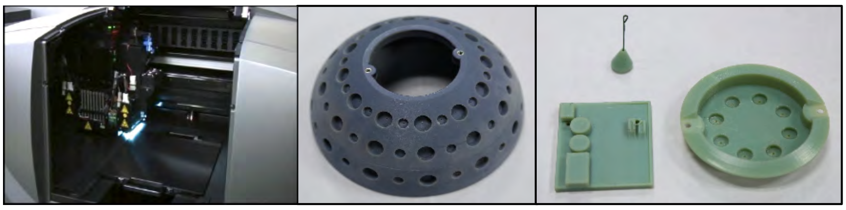


Figure 5. 3D printer with 3D printed dome, cap, towers, and jig

The synthetic tissue layers were created using Smooth-On platinum cure silicone products. These are two part silicones, which can be colored and mixed with other additives to achieve the desired product attributes such as durometer. The silicone used for the “fat” layer gave a gel-like and slightly sticky texture (Eco-flex Gel), while the “skin” silicone had a more firm and non-sticky quality (Ecoflex-0030). These silicones were chosen because they gave the closest resemblance to actual tissue properties. The fat silicone was poured directly onto the dome to the desired thickness. A clay mold was then made to replicate that thickness, which was used to form the skin layer (Figure 6). Embedded in the skin was a layer of polyester mesh, which helped to provide structure and stability of the skin. Small vessels were also created by quickly curing the silicone to a small tube. Using these materials we were able to create a set of synthetic tissues for less than $20.

Figure 6. Pouring of silicones and first HFP

The puzzle piece shape and the other markers were drawn on the skin surface using a permanent marker. The exercises were drawn on in different locations, sizes, and orientations for the first HFP. After testing the HFP on the robotic system, we finalized the size and orientation of the exercises on this new dome. This is important because as learned in the LFP stage, the exercises needed to be placed strategically to compensate for the range of movement of the robotic arms. Despite having 7 degree-of-freedom instruments, there are still limitations to the amplitude of the movement of the robotic arms. We also determined that three trials of each exercise could fit on one dome, so each work station (i.e. group of exercises) repeated at 120 degree increments on the dome. Eventually, we determined that after dissecting the three vessels significant space was available for more dissection in the fat layer. So, we added three additional vessels located to the right of the original vessels and out of range of potential damage from other exercises (Figure 7). By doing so, the fat could be used six times and the skin used three times, which incurs lower costs for the materials used during training.

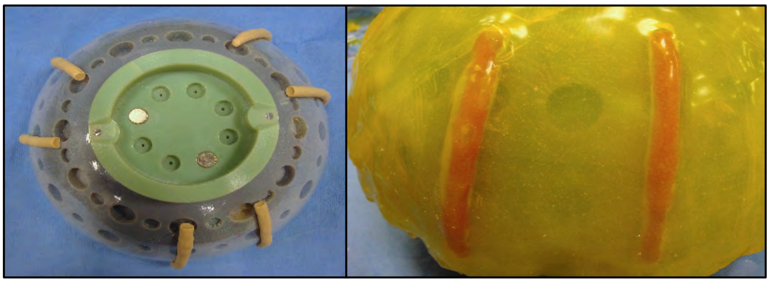


Figure 7. Vessel placement on dome and in fat

Over many iterative models, we improved our techniques and experimented with different materials and additives to achieve the desired qualities. For example we began adding a Thixotropic additive to thicken the mixture and allow us to cast the material onto a curved surface. We also tested different inks and techniques of printing the shapes and markers on the skin; however, most inks and paints cannot be used on silicone. We decided to use a silicone based paint product, which cured the design to the silicone surface.

We 3D printed miniature dome models (2” in diameter) to begin testing molding materials. We created silicone molds and used a urethane plastic to cast the model. By doing this we realized that the original 3D printed material was porous and caused bubbling in the molding, leading to surface bubbles on casted models. So, a new full sized dome was printed in a smoother and less porous material, which would be better for manufacturing. The new dome shell and cap was designed with divots only at the locations necessary for holding a tower (Figure 8).

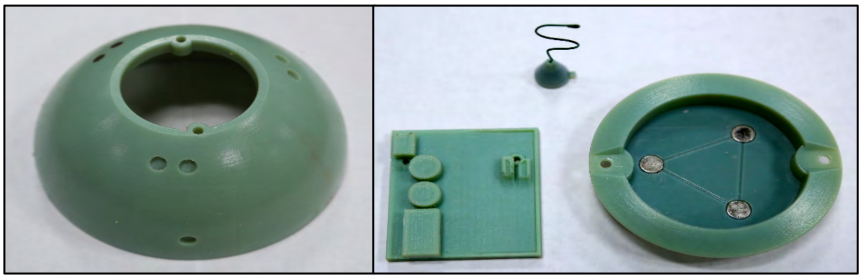
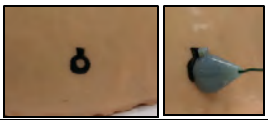
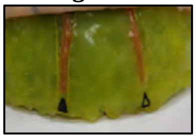


Figure 8. Final 3D printed dome shell

Since this device will be used for training and education, a high level of standardization is necessary. For this we added small markers that ensure the pieces are assembled correctly and in a standardized manner for all participants. Table 3 details the standardization pieces.

Table 3. Description of the Standardization Markers

Standardization Markers	
Tower tongues 	Used to orient the towers in the correct direction for each exercise.
Triangle in lid 	Used to show proper orientation of the towers that are placed in the cap. The towers are placed in the two locations directly in line with the puzzle piece and with the tower tongues on the corresponding line of the triangle. This ensures that the S-shaped towers face the correct direction for all users.
Tower orientation markers 	These markers are used to show the placement of the towers on the skin and the orientation of the tower. The towers are placed on the marker with the tongue aligned with the tongue mark. This ensures that all towers face the correct way.

Triangles on dome shell 	These small markers are located at 120 degree increments on the lower edge of the dome. They signify where the embedded vessels should be located when the tissue layers are placed on the shell.
Triangles on fat 	There are two types of triangle markers on the fat: open and closed. The closed triangles indicate the location of the first use vessels. When the fat is placed on the dome, the closed triangle is aligned with the triangle marker on the dome shell. After all three vessels are used, the fat is rotated and the open triangles are aligned with the triangles on the dome. This ensures that the vessels are in the accurate location for the dissection exercises.
Triangles on skin 	The triangle markers on the skin are aligned with the triangles on the fat layer. These ensure that the puzzle piece lies directly over the vessel and that the tower markers align with the underlying magnets.
Cap placement notch 	The notch in the cap ensures that users place the cap in the correct orientation. Since the magnet divots are placed in the shape of a triangle, the cap has to be secured in a specific orientation for the magnet divots to align properly.

In the final HFP, the exercises existed as they would in the manufacturing phase. Final testing was performed in order to ensure that all specifications were correct and to build a specifications document, which was used to create final CGI and CAD files (Figure 9).

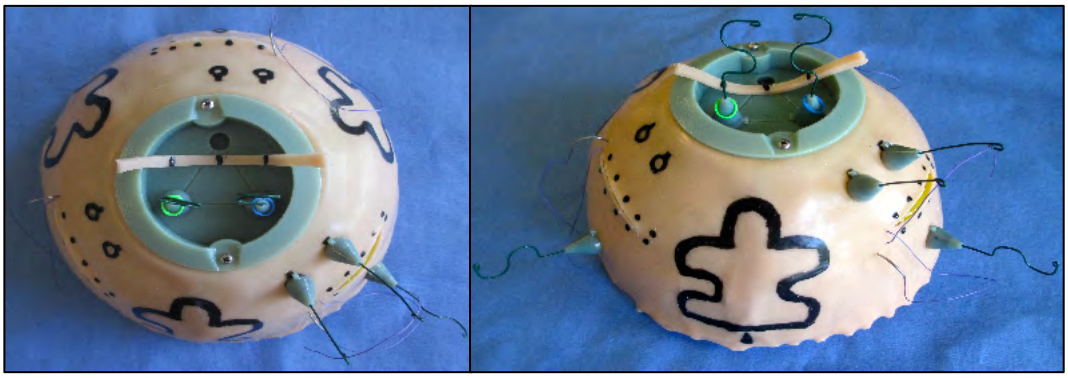


Figure 9. Final HFP

PRODUCTION

The final CGI, CAD, and specification document were sent to the manufacturing company and simulation companies to assist them in their development of physical and virtual domes (Figure 10).

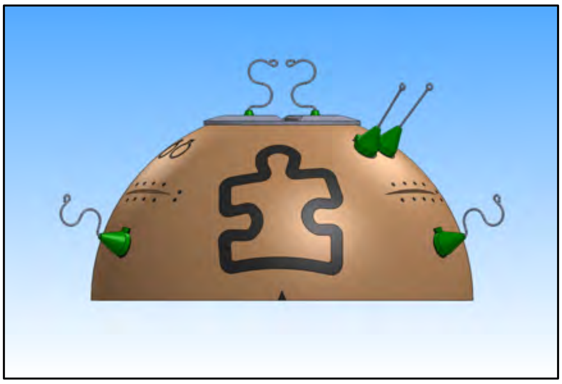


Figure 10. Final CGI

A local manufacturer, familiar with the materials used during prototype testing, used the dome and performed all of the exercises prior to beginning the process. This provided a first-hand experience of why certain material qualities were so important. The goals for this phase, in addition to mass production, were to maintain device integrity and minimize cost. Some of the materials used during prototyping were more expensive than what would be feasible for training centers. For example the $1,000 materials cost for the 3D printed dome was reduced to less than $25.

The simulation exercises of the FRS dome will be incorporated into two simulators: the da Vinci Skills Simulator (dVSS) and the Mimic dV-Trainer (Figure 11). Both systems contain the six FRS exercises, but vary in their software and hardware. The dVSS is a simulation system, which integrates with the actual console of the surgical system. This allows users to train using the exact hardware that they use when operating. The dV-Trainer is a standalone system that uses custom hardware and software. These simulations give the users experience performing the FRS exercises without requiring the use of the entire robotic surgical system. Generally, the systems are dedicated resources to the hospital surgical department and difficult to reserve for training purposes. The simulators also allow unlimited practice sessions without consuming the physical materials of the dome. The research team worked with each of the simulator companies to create and test multiple prototype versions of the exercise software. Our extensive experience with the real materials and our surgeons' experience with human surgery allowed us to critically evaluate the simulated behaviors of materials and the scoring methods. This feedback has led to significant improvements in the accuracy and usability of the simulators.

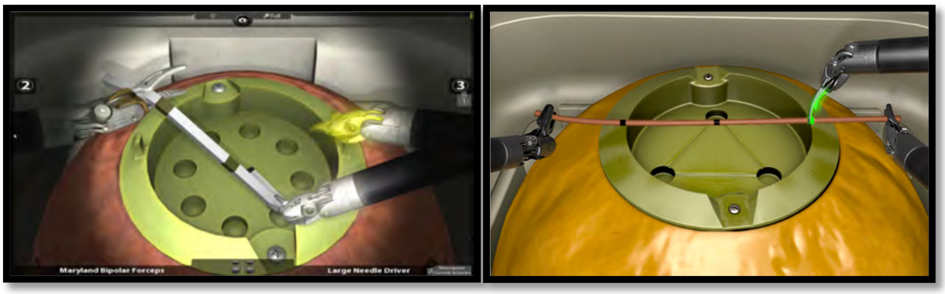


Figure 11. Mimic dV-Trainer and Simbionix's dVSS simulated dome exercises

Maintaining the simulated physical properties of the dome was paramount. Since the simulations may be used without proctors, the physical behaviors have a considerable impact on the scoring metrics and guidance that is given for improving performance. The research team evaluated the simulated exercise properties including elasticity of materials, flexibility of sutures, simulated gravity, and the effects of excess force on the virtual device to ensure that it behaved similar to the real dome. The real materials however were also limiting to some of the desired qualities, particularly in the vessel dissection exercise. The silicon-based materials act as insulators, preventing cauterization of the small vessel. Both simulators allow the user to apply energy for cauterization, as well as receiving a visual indication that the vessel is losing blood, prompting the user to manage the situation appropriately.

Some of the metrics also varied between the physical and simulated domes. While the physical dome is scored via expert video reviewing, the simulator can more objectively assess a user's performance. This allows the simulated exercises to score some errors more accurately, such as instruments being out of view for a specific amount of time and over a specific distance.

The research team will include these simulations in a pilot study and provide the simulation companies further formative feedback on the usability of their systems, to mitigate complications that may occur during the larger multi-site validation study that will follow. This pilot study will also establish preliminary scoring benchmarks based on expert performance, which will be used to guide the multi-site validation study.

CONCLUSION

Over the course of two years, we created an easily integrated device, using low cost but high-quality materials. This paper outlines the steps of the FRS dome from idea conception to the development of physical and virtual devices. The goal of this paper is to share the evolution and process for others interested in training and assessment devices. Since the FRS dome specifications are open-source, this also serves as an important resource for potential producers.

We have taken away several lessons from our experimentation that made our process a success including having a multidisciplinary team, soliciting frequent feedback, using easily adaptable designs, testing on small models, and using commercial materials during prototyping. Our multidisciplinary team of surgeons and engineers allowed for a diverse perspective during the construction of the device. The design changed many times and it was beneficial to start off using basic models that accommodated the varying designs. It was advantageous to work with actual manufacturing materials once we developed a functional prototype to better envision the final product and allow a smoother transition to the manufacturing phase. We recommend testing materials on small models, which will help cut time and costs. Finally if possible, work closely with the manufacturing teams at an early stage of development, particularly when working with virtual models. This will help to flesh out details and encourage collaborative development earlier in the process.

The next step of this work is to conduct formal validation testing of the curriculum including the device and related simulations via a pilot and national multi-site validation study. The FRS dome features basic robotic surgical skill exercises, which are applicable to most specialties. This basic device is scalable and will be the foundation for the future, more specialized FRxS devices (e.g., the Fundamentals of Robotic Gynecologic Surgery (FRGS) and the Fundamentals of Robotic Urologic Surgery (FRUS)).

REFERENCES

Barbash, G. I., Friedman, B., Glied, S. A., & Steiner, C. A. (2014). Factors associated with adoption of robotic surgical technology in US hospitals and relationship to radical prostatectomy procedure volume. *Annals of Surgery*, *259*(1), 1–6. doi:10.1097/SLA.0b013e3182a5c8b8

Blavier, A., Gaudissart, Q., Cadière, G.-B., & Nyssen, A.-S. (2007). Comparison of learning curves and skill transfer between classical and robotic laparoscopy according to the viewing conditions: implications for training. *American Journal of Surgery*, *194*(1), 115–121. doi:10.1016/j.amjsurg.2006.10.014

Chitwood, W. R., Jr, Nifong, L. W., Chapman, W. H., Felger, J. E., Bailey, B. M., Ballint, T., … Albrecht, R. A. (2001). Robotic surgical training in an academic institution. *Annals of Surgery*, *234*(4), 475–484; discussion 484–486.

Chowriappa, A., Raza, S. J., Fazili, A., Field, E., Malito, C., Samarasekera, D., … Guru, K. A. (2014). Augmented Reality Based Skill Training for Robot-Assisted Urethro-Vesical Anastamosis: A Multi-Institutional Randomized Controlled Trial. *BJU International*. doi:10.1111/bju.12704

Geller, E. J., Schuler, K. M., & Boggess, J. F. (2011). Robotic surgical training program in gynecology: how to train residents and fellows. *Journal of Minimally Invasive Gynecology*, *18*(2), 224–229. doi:10.1016/j.jmig.2010.11.003

Grover, S., Tan, G. Y., Srivastava, A., Leung, R. A., & Tewari, A. K. (2010). Residency training program paradigms for teaching robotic surgical skills to urology residents. *Current Urology Reports*, *11*(2), 87–92. doi:10.1007/s11934-010-0093-9

Hubens, G., Coveliers, H., Balliu, L., Ruppert, M., & Vaneerdeweg, W. (2003). A performance study comparing manual and robotically assisted laparoscopic surgery using the da Vinci system. *Surgical Endoscopy*, *17*(10), 1595–1599. doi:10.1007/s00464-002-9248-1

Jarc, A. M., & Curet, M. (2014). Construct validity of nine new inanimate exercises for robotic surgeon training using a standardized setup. *Surgical Endoscopy*, *28*(2), 648–656. doi:10.1007/s00464-013-3224-9

Levy, J. (n.d.). Fundamentals of Robotic Surgery: Online curriculum. Retrieved from http://frs.casenetwork.com/lms/

Luca, F., Valvo, M., Ghezzi, T. L., Zuccaro, M., Cenciarelli, S., Trovato, C., … Biffi, R. (2013). Impact of robotic surgery on sexual and urinary functions after fully robotic nerve-sparing total mesorectal excision for rectal cancer. *Annals of Surgery*, *257*(4), 672–678. doi:10.1097/SLA.0b013e318269d03b

Maan, Z. N., Gibbins, N., Al-Jabri, T., & D'Souza, A. R. (2012). The use of robotics in otolaryngology-head and neck surgery: a systematic review. *American Journal of Otolaryngology*, *33*(1), 137–146. doi:10.1016/j.amjoto.2011.04.003

Patel, H. D., Mullins, J. K., Pierorazio, P. M., Jayram, G., Cohen, J. E., Matlaga, B. R., & Allaf, M. E. (2013). Trends in renal surgery: robotic technology is associated with increased use of partial nephrectomy. *The Journal of Urology*, *189*(4), 1229–1235. doi:10.1016/j.juro.2012.10.024

Serati, M., Bogani, G., Sorice, P., Braga, A., Torella, M., Salvatore, S., … Ghezzi, F. (2014). Robot-assisted Sacrocolpopexy for Pelvic Organ Prolapse: A Systematic Review and Meta-analysis of Comparative Studies. *European Urology*. doi:10.1016/j.eururo.2014.02.053

Smith R, Patel V, Satava R. (2013). Fundamentals of robotic surgery: a course of basic robotic surgery skills based upon a 14-society consensus template of outcomes measures and curriculum development. Int J Med Robot Comput Assist Surg MRCAS.

Smith, R., Patel, V. R., Chauhan, S., & Satava, R. (2013, February 20). *Fundamentals of Robotic Surgery: Outcomes Measures and Curriculum Development*. NextMed/ MMVR, San Diego, CA.

Zorn, K. C., Gautam, G., Shalhav, A. L., Clayman, R. V., Ahlering, T. E., Albala, D. M., … Members of the Society of Urologic Robotic Surgeons. (2009). Training, credentialing, proctoring and medicolegal risks of robotic urological surgery: recommendations of the society of urologic robotic surgeons. *The Journal of Urology*, *182*(3), 1126–1132. doi:10.1016/j.juro.2009.05.042

Zureikat, A. H., Moser, A. J., Boone, B. A., Bartlett, D. L., Zenati, M., & Zeh, H. J., 3rd. (2013). 250 robotic pancreatic resections: safety and feasibility. *Annals of Surgery*, *258*(4), 554–559; discussion 559–562. doi:10.1097/SLA.0b013e3182a4e87c

Robotic Simulators: A Case for Return on Investment

Roger D. Smith PhD
Florida Hospital Nicholson Center
Celebration, FL 34747
roger.smith@flhosp.org

Khara M. Simpson MD
Columbia University Medical Center
New York, NY 10032
kmsimpmd@yahoo.com

ABSTRACT

Simulation has been integrated into the education and certification process in aviation and military arenas with significant success in providing cost effective training. The transition from the apprenticeship model to simulation has been slower in the field of medicine with cost, lack of curricula and high fidelity exercises and equipment being the main reasons. With recent improvements in all areas, cost remains a significant challenge.

This report describes our novel analysis of the return on investment (ROI) that can be achieved through the inclusion of simulator use within a robotic surgery business practice and as an alternative source of training revenue. Information was gathered through an extensive literature review and expert interviews for the development of an interactive calculator for institutions to utilize when considering an investment in robotic surgery simulators.

This ROI model presents the core improvements to existing operations which may be realized through the use of simulators of robotic surgery. Category headings include simulator investment costs, surgeon productivity, surgeon health, hospital costs, and other training costs. The user of the model is able to enter their own numbers for their unique facilities. The spreadsheet model will calculate the costs and benefits associated with each area, create category subtotals, and then an overall total for all areas. Using these numbers, it can then calculate an ROI percentage for the simulators. This model represents one tool to assist organizations in making the investment in these devices and training programs.

ABOUT THE AUTHORS

Roger Smith, PhD, is an expert in the development of simulation devices and training programs. He has spent 25 years creating leading edge simulators for the Department of Defense and Intelligence agencies, as well as accredited methods for training with these devices. He is currently the Chief Technology Officer for the Florida Hospital Nicholson Center where he is responsible for establishing the technology strategy and leading technology implementation. He has served as the CTO for the U.S. Army PEO for Simulation, Training and Instrumentation (PEO-STRI); VP and CTO for training systems at Titan Corp; and Vice President of Technology at BTG Inc. He holds a Ph.D. in Computer Science, a Doctorate in Management, and an M.S. in Statistics. He has published 3 professional textbooks on simulation, 10 book chapters, and over 100 journal and conference papers. His most recent book is *Innovation for Innovators: Leadership in a Changing World.* He has served on the editorial boards of the *Transactions on Modeling and Computer Simulation* and the *Research Technology Management* journals.

Khara Simpson, MD is a second year fellow in minimally invasive surgery at Columbia University, where she serves as an assistant attending and instructor of obstetrics and gynecology. She completed her medical school education at Howard University College of Medicine where she was inducted into the Alpha Omega Alpha honor medical society. Following, she completed her OB/GYN residency at Johns Hopkins University and served as administrative chief resident. She recently completed a one year research fellowship at the Florida Hospital Nicholson Center focusing on robotic surgery simulation. Her additional research interests include resident education and simulation training, and best practices to promote cost effective care.

Robotic Simulators: A Case for Return on Investment

Roger D. Smith PhD
Florida Hospital Nicholson Center
Celebration, FL 34747
roger.smith@flhosp.org

Khara M. Simpson MD
Columbia University Medical Center
New York, NY 10032
kmsimpmd@yahoo.com

INTRODUCTION

Creating a viable robotic surgery practice within a hospital is an expensive and risky endeavor. The investment in equipment, facilities, personnel, and process modification is significant, amounting to millions of dollars in the early years of a program. Many hospitals make this investment with a limited understanding of how best to structure a robotics practice and the probability of achieving a positive return on this investment. At the business end of a decision to create a robotic surgery practice, the hospital has the goals of optimizing the utilization of the robotic operating room, reducing costs, and ensuring patient safety. Training to and maintaining the competency of the surgeons performing the procedures has a direct impact on these areas, allowing the inclusion of simulators in the robotic business unit to make valuable contributions. The proposed effects are summarized in Figure 1. This report describes our analysis of the return on investment that can be achieved through the inclusion of simulators and their regular use within a robotic business practice; and as an alternative source of training revenue.

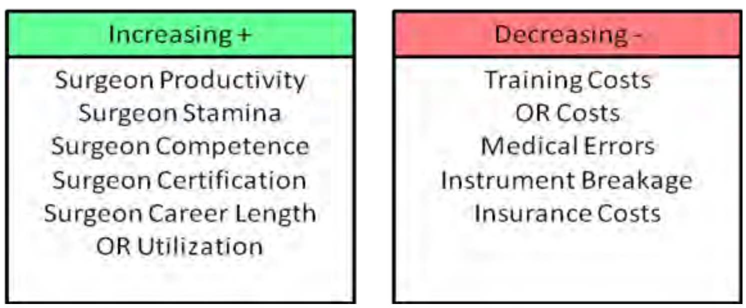


Figure 1. Summary of Simulation Effects on Surgical Practice

BACKGROUND

For every complex and expensive system there emerges a need for training devices and scenarios that will assist new learners in mastering the use of the device and understanding how to apply it with value. Intuitive Surgical's da Vinci robot is just such a system. It is currently the only FDA approved device for laparoscopic robotic surgery on human patients. Despite the 1.5 to 2 million dollar price tag, the device has seen rapid distribution and its implementation has led to the need to develop more efficient and effective training methods, as well as assessment and skill maintenance tools. In laparoscopic surgery, simulators have played an important role in improving the practice of surgery over the last 20 years (Schout and Hendriks, 2010; Wohaibi and Bush, 2010). The same trends and values will likely apply to robotic surgery.

The complexity, criticality, and cost associated with the application of the da Vinci surgical robot have stimulated the commercial creation of simulators which replicate the operations of this robot. There are currently three different simulation systems available for training and developing skills in robotic surgery: da Vinci Skills Simulator (Intuitive Surgical Inc.); dV-Trainer (Mimic Technologies, Inc.); and RoSS (Simulated Surgical Skills LLC). Each of these possesses unique traits which make them valuable solutions for different types of users and learning environments. Investment in the simulators alone represents a major capital investment as costs range from $100K (RoSS/dV-Trainer) to $600K (da Vinci Skills Simulator) per device and its associated support equipment. Coupled with increases in direct costs from operating room time and supplies, (Venkat and Chen, 2012; Pasic and Rizzo, 2010; Bolenz and Gupta, 2010; Barnett and Judd, 2010; Holtz and Miroshnichenko, 2010) robotic surgery implementation can have a profound impact on hospital finances. This doesn't include costs associated with ensuring patient safety and surgeon training. From a business perspective, a hospital, college of medicine, or robotic practice should identify a return that will be achieved with the purchase of these devices.

History of Simulation

Rehearsal and simulation has been one of the primary means of developing and maintaining proficiency in specific skills for thousands of years. Lectures and written materials are the primary means for developing cognitive knowledge, but these are not effective at instilling psychomotor skills in any field. The skills needed in the hands, body, and coordination with the mind must be developed through practice. Rehearsal in a non-lethal environment has become the standard of practice for learning in military warfare and aviation. The lives of soldiers and pilots today are considered of significant value to demand a structured training program and measure of proficiency before entering a life threatening situation. Given the technology that has emerged in the late 20th and early 21st centuries, the means of rehearsal have shifted from drills toward events that are mediated by computers with the ability to measure performance and provide constructive feedback on means to improve that performance. Implementation of direct practice on simulators has led to significant quality improvements and a reduction in training costs to the tune of hundreds of millions of dollars. There is a 30 to 40 year history of success documented in the military literature, specifically regarding its return on investment. On average simulators cost 5-20% of live training and can reduce the length of training by 10-30% (Fletcher and Alexander 2013). Simulation has also been found to be well liked by trainees and found to be comparable to live training in terms of experience as well as outcome (Worley and Simpson 1996).

Simulation devices are similarly a prominent tool in medical and surgical education. From the most primitive stuffed dolls with anatomical markings used in ancient Chinese societies, to the most current computer driven, three dimensional representations of living tissue, simulators have been helping to train surgeons for over a thousand years . While there are parallels between the military and medical fields regarding the involvement of human life and the need for structured high fidelity training, there are some unique challenges. There are many confounders in medicine and the seemingly direct relationship between simulation and improved patient outcomes and reduction in costs is not robust. So, while most agree that simulation is a necessary requirement for surgical training, the lack of evidence coupled with the costs of device and curriculum development have led to a slower adoption of simulation in surgical education when compared to industry. One of our goals is to use the lessons from our military and aviation industries in conjunction with published literature to identify key variables that impact the ROI of robotic surgical simulators and provide a loose framework for how simulators can be better utilized to achieve financial, educational, and patient safety benefits. We will focus on the concept of return on investment (ROI) for these simulators, not on the capabilities of one specific device. Our goal is to assist the purchasers and users of these devices in determining whether such a device is a sound financial investment.

Published Literature

There are very few studies evaluating ROI for the simulators of the robot. There is one article directly evaluating the RoSS simulator (Rehman and Raza 2013) where annual training hours were converted to training time on a robotic console. In that study, the use of the stand-alone simulator resulted in cost savings of $600,000. Animate lab training of the same duration would have cost approximately $72,000 annually to train 100 people each year. The remainder of the studies look solely at laparoscopic simulators. For example, in March 2004, Frost & Sullivan Inc. conducted a ROI study on three training simulators sold by Immersion Medical. Specifically, the data regarding the Laparoscopy AccuTouch System was published. The Laparoscopy AccuTouch System uses advanced 3D technology and graphics to re-create the procedures and environment of abdominal laparoscopic surgery. Using reported median values from survey data, financial benefits were estimated at $168,767, based on annual costs $76,000, with an estimated payback period of 169 days. Also somewhat related, is a review article by Leddy et al that reviewed the published cost analyses of robotic surgery for both urology and pediatric surgery, in an effort to develop a novel model for determining return on investment. The premise for the model was that reduction in costs directly relate to the ability of the technology to reduce hospital length of stay. Surgical volume can be limited by the availability of hospital beds, in which shorter hospital stays lead to greater bed availability, which can lead to more procedures being performed. It also acknowledged the learning curve of robotic surgery and the expectation for OR times to decrease with improved technology, surgical technique, and time.

METHODS

The first step was to identify, at a basic level, the main effects that robotic simulation would have on a healthcare organization. We found that the use of a simulator can have a direct impact on at least three major, but separate parts of the healthcare delivery process (Figure 2).

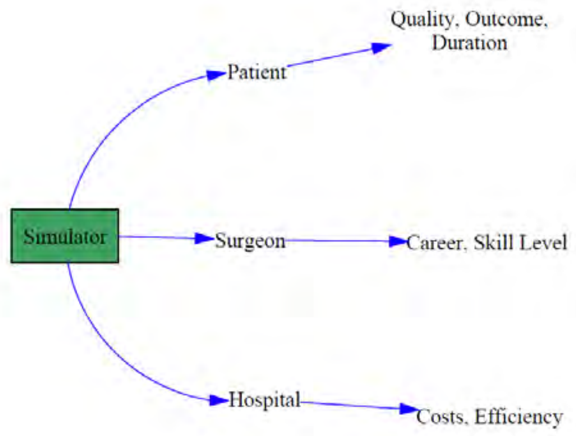


Figure 2. Robotic Surgery Simulator Impact Areas

The first, and most important, is patient outcome. The use of a simulator can potentially improve patient outcome by allowing the surgeon to perform more effectively and efficiently with fewer adverse events.

The second is the impact on the surgeon. He or she may find that the ability to use the robot effectively impacts their career with additional options that can be offered to the patient and to a potential hospital employer. It may also reduce the long-term wear and tear on the surgeon's body due to improved ergonomics, making it possible to continue practicing surgery for a longer period of time.

Third is the effect that it will have on the costs and efficiency of the business practice. Simulator trained surgeons may perform procedures more rapidly, with less error, and lower instances of equipment breakage. This ties directly back to improved patient outcomes.

With this as a foundation, we used a diagramming method popularized by Jay Forrester at Massachusetts Institute of Technology (MIT) to explore the interrelationships between a large number of variables within the complex and dynamic hospital surgical program. This method identified variables which react almost immediately to changes in training offered to robotic surgeons as well as longer-term variables which may take years to emerge as measurable returns on the investment. These relationships can be grouped and organized as shown in Figure 3. Note that each variable is associated with a calculable financial outcome. This model was useful in identifying the factors which contributed to ROI in a first, second, and third level of separation. It also motivated discussions around the financial and non-financial variables that were changing with the incorporation of simulation.

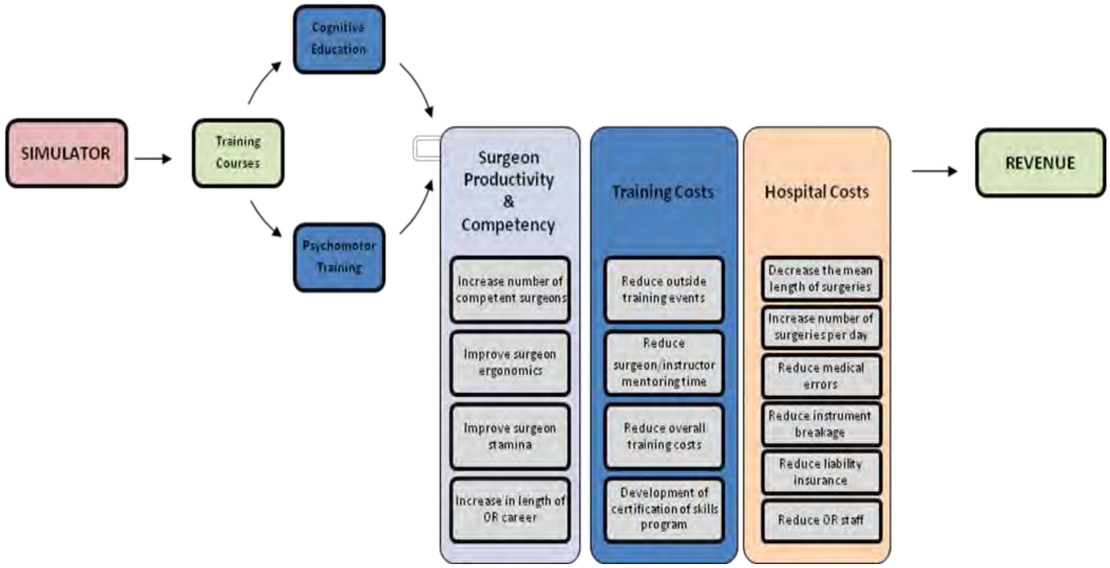


Figure 3. Effects of Simulation-based Training on Robotic Surgery Business

Expert Opinion

Twelve expert interviews were then conducted at several simulation centers around the country and internationally to further expound upon the model, including identifying and quantifying variables. Surgical leaders interviewed included Randy Fagin MD, Texas Institute of Robotics; Brendan Sayers PA, University of Texas – Southwestern; Arnold Advincula MD, Celebration Health; Thomas Lendvay, MD, University of Washington College of Medicine; Robert Sweet, MD, University of Minnesota College of Medicine; Col Timothy Brand, MD, Madigan Army Medical Center; Jacques Hubert, MD, University of Lorraine Medical School; Dimitrios Stefanidis, MD, Carolinas Healthcare System; Martin Martino, MD, Lehigh Valley Health Network; Mona Orady, MD, Cleveland Clinic; John Lenihan, MD, MultiCare Health Systems; and Michael Pitter, MD, Newark Beth Israel Medical.

There were several common themes identified from the expert interviews. Most agreed that having a defined curriculum was imperative to successful training and that it should be identified prior to purchasing a simulator. These robotic training curriculums consisted mostly of manufacturer-developed didactic education with a combination of simulator and/or dry lab assessment. It was also clear that there were a variety of successful arrangements for location and proctoring of a robotic simulation curriculum. The majority utilized a combination of the dVSS and the dV-Trainer. Only one program had a dedicated training console, therefore most simulation with the dVSS was limited to nights and weekends. Simulation was most often a part of resident and fellow curriculums but approximately 50% of the sites also had attending level programming. Simulator value was also identified to be tied directly to the accessibility of the necessary equipment.

They verbalized a strong need for studies evaluating correlations between simulation and improved clinical outcomes, determining the ideal length and intervals for simulation training, and comparative studies between institutions with and without simulation centers. In addition to maintenance and certification, curricula could also be used for remediation training following complications. Software interests surrounded the development of procedure specific training and increasing complexity of these procedures. Several believed that simulation could be used globally to decrease insurance costs and increase reimbursements.

Simulator ROI Model

Using the information in the system dynamics diagram, we created a basic model of the ROI with short-term, measurable returns. The model is populated with average numbers from the literature and the authors' institution, but the user is able to enter specific numbers representing their unique facilities. Users enter data for their specific institution and the model calculates intermediate values and an annual return for Years 1 & 2 following the investment. Interested readers may request the full calculator with relevant appendices from the authors at no charge. Figure 4 illustrates part of the data entry and calculation fields for the calculator. Table 1 provides a sample of the variables and returns that are possible from a simulator-based training program.

Robotic Simulator Return on Investment Model				
	User Entered Data	*Orange Fields*		
	Model Calculated Data	*Grey Fields*		
			Overall Average	
		Initial Investment	Annual	Totals
Simulator Investment Costs				
	Cost of Simulator	$100,000	$0	
	Number Purchased	4	4	
	Annual Maintenance Fee	$0	$10,000	
	Simulator Investment	$400,000	$40,000	
	Annual Maintenance Fee	$0	$10,000	
	Facility Costs	$100,000	$5,000	
	Staffing Costs	$20,000	$50,000	
	Supplies & Materials	$1,000	$200	
	Investment	$521,000	$105,200	$105,200
Surgeon Productivity				
	Number of Surgeons Trained		2	
	Mean Length of Operation		2	hours
	Time Saved per Procedure		20	minutes
	Length of Surgical Day		8	hours
	Surgeries per day		8	
	New Free Surgery Time per day		2.67	hours
	Added Surgeries per Day		1	
	Revenue per Surgery		$32,000	
	Additional Revenue per day		$32,000	
	Surgical Days		100	per year
	Additional Revenue per Year		$3,200,000	$3,200,000
Surgeon Health				

Figure 4. ROI Calculator (Partial Screenshot)

Table 1. ROI Calculator Sample Variables

INVESTMENT	Cost (Year 1)	RETURNS		Benefit (Year 1)
Cost of Simulator	**$100,000**	*Surgeon Productivity*	Number of Surgeons Trained	**2**
Annual Maintenance Fee	**$0**		Mean Length of Operation (Hours)	**2**
Facility Costs	**$100,000**		Time Saved per Procedure (Minutes)	**20**
Staffing Costs	**$20,000**		Length of Surgical Day	**8**
Supplies & Materials	**$1000**		Revenue Per Surgery (Includes Admission)	**$32,000**
			Surgical Days per Year	**100**
		Surgeon Health	OR Stamina (Hours)	**6**
			Improved Stamina (Hours)	**1**
		Hospital Costs	Liability Insurance	**$100,000**
			Competence Discount	**5%**
			Average Instrument Breakage	**$5,000**
			Breakage Reduction Factor	**10%**
			Surgical Error Rate	**1%**
			Error Improvement Rate	**10%**
		Training Costs	OR Training Time per Day (Hours)	**1**
			Number of In-house Courses	**6**
			External Training Event Cost (w/Travel)	**$6,000**
			Number of External Events	**1**
			Training Reduction	**50%**

Listed below are the many factors to be considered in each cost category as well as limitations to the calculator.

Investments

The costs associated with adopting a simulator training program include:

Cost of Simulator. The initial investment in purchasing the simulator device. The price point for many surgical simulators is in the neighborhood of $100,000 per device.

Annual Maintenance Fee. Most devices include one year of maintenance or service warranty with the purchase of the device. In following years, there is an annual fee to cover the installation of updates and repairs to the device. One must also allow for additional repairs beyond the normal maintenance.

Facilities Costs. The equipment needs to be housed in a location that is conducive to sufficient access and training. A dedicated space is preferred, but some institutions use a shared space model to reduce this cost. This impacts the degree to which surgeons can use the devices on their own initiative which directly impacts training efficiency.

Staffing Costs. Most simulator devices require a knowledgeable trainer to instruct the surgeons on the use of the device and to facilitate data collection and the tracking of performance improvement. This staff person does not necessarily have to be dedicated to the robotic simulators, but can be a shared resource that has additional duties in the hospital system. The staffing cost can vary widely depending on the type of assignment used, such as hiring new personnel versus training an existing employee. An additional consideration would be the loss of revenue from physicians, nurses, and other medical professionals during training sessions.

Supplies & Materials. Because robotic simulators are electronic and use no real consumables like suture pads, needles, and fluids, the supplies and materials for continuing to use them is minimal. Some typical supplies include: power strips, extension cords, cleaning wipes, spare parts, computer thumb drives, and computer adapters.

The investment in simulator programs is typically highest at initial purchase for the equipment and much lower for each ensuing year. Therefore, the model separates the first year from years following. The expenses in years 2, 3, 4, etc. appear to be very similar. At this point there is little data on equipment fatigue with use. There is only published data on the RoSS. Master controllers and the pinch devices need to be replaced at 180 and 360 hours of use respectively (Rehman 2013). We are therefore not able to identify the effective useful lifetime of a simulator device or a significant upturn in costs due to increased maintenance in later years. Insight into what to expect in robotics may be found in the operating expenses for laparoscopic training simulators which have been in use for many more years.

Returns

The returns that can be experienced through the implementation of a robotic simulator training program fall into four major categories: surgeon productivity, surgeon health, hospital costs, and training costs. Each of these has many variables which contribute to financial returns.

Surgeon Productivity. The primary goal for most organizations creating a simulator training program is to increase the productivity of their surgeons. A model of surgeon productivity begins with the number of surgeons who are trained and the amount by which this training can speed up a typical operation. Reducing the time required to perform a surgical procedure can contribute to multiple variables. First, it may reduce staffing costs by allowing the staff that supports the surgeon to move on to other activities, or to work fewer paid hours during the day. In some cases, it may reduce total OR staffing due to the ability of the surgeon to control both the camera and the surgical arms concurrently. Second and most importantly, if the reduction is large enough, it may open a window in the OR schedule which is sufficient to perform an additional operation during the day. Third, if simulator-based training can reduce the variance in surgical times for procedures, it will make the scheduling of procedures more accurate. This would reduce the "slippage" that occurs in the daily schedule because a procedure takes twice as long as was scheduled. Smaller variances lead to more efficient scheduling which can reduce staffing and facilities costs. This model focuses on the impact of reducing surgical time sufficiently to open a window for an additional surgery. The impacts of variance across multiple instances of similar procedures is worthy of an entire dedicated study and is not included.

Surgeon Health. Robotic systems significantly reduce the physical workload experienced by traditional open and MIS surgeons. All surgeons have a level of stamina that allows them to perform for a specific number of hours each day. Upon reaching this point, the muscles and mental focus of the surgeon are diminished. Continuing to operate can bring risks both to the health of the patient and the surgeon.

Rehearsal in a simulator has the effect of training the muscles which are used in surgery. Familiarization with the tasks also reduces the amount of mental and physical effort that is required to complete each procedure. The extension of surgeon stamina and the reduction in energy expended can both lengthen the number of hours that a surgeon is able to perform optimally.

This physical and mental training can extend the surgical day for a specific individual. It may also extend the length of the surgeon's operating career by reducing repetitive stress injuries and aging due to mental stress. The cumulative impact on the length of careers cannot be calculated at this point. Exploration of this effect is left for future researchers.

The model of surgeon health included in the ROI model in this paper is limited to daily extensions of operating time.

Hospital Costs. The model identifies four major hospital cost categories which may be reduced due to improved surgeon competence: instrument breakage, liability insurance, surgical error rate, and training costs.

Inexperienced surgeons can break the controls of the robot through the exertion of too much force and fighting against the electrical and mechanical components of the robot. When the surgeon damages the robot, the cost of many (but not all) repairs is covered by the maintenance agreement; but the significant cost to the hospital is in the disruption that occurs in postponing or rescheduling an operation because the robot is no longer operable. Inexperienced surgeons can also damage the surgical instruments held by the robot. Each of these typically costs between $1,200 and $2,500. Frequent abuses come from forceful instrument collision and friction along the shafts. Simulator training can develop skill and dexterity with an instrument which prevents this from happening.

As medicine becomes even more evidence based, there will be the opportunity to quantify the liabilities associated with specific surgeon competence levels. This may allow surgeons who can document additional training and acquired competence to reduce their medical liability costs when compared with surgeons who have less training. Experts whom we interviewed for this study indicated that they had already begun annual mandatory simulation-based exams as part of their internal risk management plan with the hope of realizing insurance savings in the future.

The competence that comes from simulator-based training can reduce the number of errors that are made in robotic surgery. Errors can require a return to the operating room for a procedure which generates no additional revenue.

There are multiple forms of training used to acquire, maintain, and extend the expertise of the surgeons. When a simulator program is introduced it can reduce the need for alternate forms of training. One popular form of robotic training today is through visitation and instruction in the OR itself. This long established and traditional practice is used in all surgical specialties. But, it is known that “in OR” training results in an extension of the time to complete the procedure (Koperna 2004). Additional time is required to explain the procedure and what is happening to the trainee. Given the high cost of OR time, a reduction in OR-based live training can generate meaningful savings for a hospital.

There are two typical forms of outsourced training. The first is bringing an instructor into the hospital system to instruct the surgeons. The second is sending the surgeons to other facilities or congress meetings to acquire the skills they need. Both of these are subject to reductions when similar skills can be acquired and measured in the simulator program. Once equipped and experienced, the organization can also become a vendor of training services as well as provide simulator rentals without the responsibility and costs of providing training staff and materials. Other organizations may seek out such rental agreements to avoid the capital purchasing costs.

Two additional considerations include team training and surgical assist training. Some of the increased length of robotic surgical cases is due to set-up time and docking times. Ensuring proper training of surgical assists and nursing in addition to physician training may decrease these times and costs.

Upper Limit

There are upper limits to the improvements that can be achieved. For actual surgical procedures, these are driven by the number of robots available to perform surgery. Increasing surgeon availability beyond the capacity of the robotic OR will not provide improvements. An additional consideration is the availability of the robotic simulator for daily use or the training efficiency of the robotic simulator. If an institution chooses to invest in a DVSS trainer but does not have a dedicated training console, then training availability will be limited when compared to those with a standalone trainer.

CONCLUSION

In 1927, William Mayo famously stated that, “There is no excuse for the surgeon to learn on the patient.” In the ensuing decades, there have been many advances in the education of surgeons. Simulators offer the next improvement along this path with capabilities that are very difficult to match through any other mode of training. We are entering a period in which competence-based learning is becoming the standard. This is an environment in which simulators possess undeniable advantages and empower the preparation of demonstrably competent surgeons.

Robotic simulators are available but are not yet a standardized part of the training process. As the airlines and the military discovered in training pilots and combat personnel, simulation devices provide so many advantages in the learning process that they have become a mandatory part of those training curricula. The expectation is that the same will occur with robotic simulators over time.

We recognize that many of the proposed financial and non-financial benefits are theoretical and that simulation in a well-designed curriculum purports different benefits than just the simulators themselves. There is an absence in the literature regarding the direct correlation between simulation and many of the positive impacts noted in this study. Despite this absence, there is evidence of three things that make these impacts real: the evidence of transfer of training with simulation as evidenced by validation studies, the presence of a steep learning curve (trends of up to 100 cases before true competency is achieved), and the improvements in surgical outcomes with high volume surgeons and centers. We also know that hospital systems and practices are highly variable and all factors cannot be identified and/or quantified financially. If there is not a direct return on investment, we are not suggesting that simulation based education is of no value. We designed the calculator to try to capture simulation effects on a business practice not only to justify the investment but to identify best practices and ways to implement robotic surgical curriculum in a meaningful and comprehensive way. These concepts should guide future research.

To summarize, the use of simulators in training robotic surgeons requires an investment in equipment, staff, facilities, and supplies. But it also offers a return on this investment in surgeon productivity, surgeon health, hospital costs, and other training costs. The realizable returns will vary by institution based on the specifics of implementation and the existing ecosystem in which they are inserted. The ROI model in this report is one tool to assist in calculating the return that can be expected by organizations that make the investment in these devices and training programs.

REFERENCES

Barnett, J. C., Judd, J. P., Wu, J. M., Scales Jr, C. D., Myers, E. R., & Havrilesky, L. J. (2010). Cost comparison among robotic, laparoscopic, and open hysterectomy for endometrial cancer. *Obstetrics & Gynecology*, *116*(3), 685-693.

Bolenz, C., Gupta, A., Hotze, T., Ho, R., Cadeddu, J. A., Roehrborn, C. G., & Lotan, Y. (2010). Cost comparison of robotic, laparoscopic, and open radical prostatectomy for prostate cancer. *European urology*, *57*(3), 453-458.

Fletcher, J. D., & Wind, A. P. (2013). Cost considerations in using simulations for medical training. *Military medicine*, *178*(10S), 37-46.

Frost & Sullivan. (2004). Return on investment study for medical simulation training: Immersion Medical, Inc. Laparoscopy AccuTouch System. Industrial research report available at: http://www.healthleadersmedia.com/content/138774.pdf

Holtz, D. O., Miroshnichenko, G., Finnegan, M. O., Chernick, M., & Dunton, C. J. (2010). Endometrial cancer surgery costs: robot vs laparoscopy. *Journal of Minimally Invasive Gynecology*, *17*(4), 500-503.

Koperna, T. (2004). How long do we need teaching in the operating room? The true costs of achieving surgical routine. *Langenbeck's Archives of Surgery*,*389*(3), 204-208.

Leddy, L. S., Lendvay, T. S., & Satava, R. M. (2010). Robotic surgery: applications and cost effectiveness. *Open Access Surgery*, *3*, 99-107.

Pasic, R. P., Rizzo, J. A., Fang, H., Ross, S., Moore, M., & Gunnarsson, C. (2010). Comparing robot-assisted with conventional laparoscopic hysterectomy: impact on cost and clinical outcomes. *Journal of minimally invasive gynecology*, *17*(6), 730-738.

Rehman, S., Raza, S. J., Stegemann, A. P., Zeeck, K., Din, R., Llewellyn, A., ... & Guru, K. A. (2013). Simulation-based robot-assisted surgical training: A health economic evaluation. *International Journal of Surgery*, *11*(9), 841-846.

Schout BM, Hendriks AJ, Scheele F et al. (2010). Validation and implementation of surgical simulators: a critical review of present, past, and future. *Journal of Surgical Endoscopy*, 24(3):536-46.

Venkat, P., Chen, L. M., Young-Lin, N., Kiet, T. K., Young, G., Amatori, D., & Chan, J. K. (2012). An economic analysis of robotic versus laparoscopic surgery for endometrial cancer: Costs, charges and reimbursements to hospitals and professionals. *Gynecologic oncology*, *125*(1), 237-240.

Wohaibi EM, Bush RW, Earle DB et al. (2010). Surgical resident performance on a virtual reality simulator correlates with operating room performance. *Journal of Surgical Research* 160(1):67-72.

Worley, D. Robert, et al. *Utility of Modeling and Simulation in the Department of Defense: Initial Data Collection*. No. IDA-D-1825. Institute for Defense Analysis, Alexandria, VA, 1996.

Comparison of the Usability of Robotic Surgery Simulators

Alyssa Tanaka, M.S., Courtney Graddy, M.S., Roger Smith, Ph.D.
Florida Hospital Nicholson Center
Celebration, FL
Alyssa.tanaka@flhosp.org, courtney.graddy@flhosp.org, roger.smith@flhosp.org

Haider M. Abdul-Muhsin, M.D.
Mayo Clinic
Scottsdale, AZ
Hma980@gmail.com

ABSTRACT

The introduction of simulation into minimally invasive robotic surgery is relatively recent and has seen rapid advancement; therefore, a need exists to develop training curriculums and to identify systems that will be most effective at improving surgical skills. Several robotic simulators have been introduced to support these aims, but their effectiveness has yet to be fully evaluated.

Currently, there are three simulators -- the daVinci Skills Simulator, Mimic dV-Trainer, and Surgical Simulated Systems' RoSS. While multiple studies have been conducted to demonstrate the validity of each system, no studies have been conducted which compare the value of these devices as tools for education and skills improvement.

This paper presents the results of an experiment comparing value, usability, and validity of all three systems. Subjects who were qualified as medical students or physicians (n=105) performed one exercise on each of the three simulators and completed two questionnaires, one regarding their experience with each device and a second regarding the comparative effects of the simulators. This data confirmed the face, content, and construct validity for the dV-Trainer and Skills Simulator. Similar validities could not be confirmed for the RoSS. Greater than 80% of the time, participants chose the Skills Simulator in terms of physical comfort, ergonomics, and overall choice. However, only 55% thought the skills simulator was worth the cost of the equipment. The dV-Trainer had the highest cost preference scores with 71% percent of respondents feeling it was worth the investment.

This work is the second component of a three-part analysis. In the previous study, the simulators were objectively reviewed and compared in terms of their system capabilities. The third part will evaluate the transfer of training effect of each simulator. Collectively, this work will offer end users and potential buyers a comparison of the value and preferences of robotic simulators.

ABOUT THE AUTHORS

Alyssa D.S. Tanaka, M.S. is a Systems Engineer at Florida Hospital's Nicholson Center. Her research work focuses on robotic surgery simulation and effective surgeon training. Her current projects include rapid prototyping of surgical education devices, the validation of a robotic surgical curriculum and evaluation of robotic simulation systems. She is a Modeling and Simulation PhD student at the University of Central Florida and previously earned a M.S. in Modeling and Simulation, Graduate Simulation Certificate in Instructional Design, and a B.S. in Psychology and Cognitive Sciences from the University of Central Florida.

Courtney Graddy, M.S. is a Human Studies Research Coordinator at the Celebration Health Research Institute where she manages projects aimed at improving patient health outcomes, employee health, process improvement and simulation research. Her current projects focus on integrating technology into standard of care and evaluating its effects on patient health and patient satisfaction, as well as evaluating teaching modalities used to train surgeons. Her career began at the North Florida South Georgia Veterans Health System where she aided in the development of employee education materials and program planning and evaluation with the Geriatric Research Education and Clinical Center. She holds a Bachelors of Science in Health Education from the University of Florida and a Masters of Health Administration from the University of South Florida.

Haidar Mohamed Abdul-Muhsin, M.D. is an endourology fellow at Mayo Clinic in Scottsdale, AZ. Prior to this role, he was a urology fellow at the Global Robotics Institute at Florida Hospital at Celebration. Dr. Abdul-Muhsin has been practicing medicine for 10 years. His research interests include prostate cancer research with emphasis on robotic prostatectomy outcomes and research related to robotic surgical simulation. Currently, he is researching the contribution of laparoscopic surgical experience in developing robotic proficiency.

Roger Smith, Ph.D. is an expert in the development of simulation devices and training programs. He has spent 25 years creating leading edge simulators for the Department of Defense and Intelligence agencies, as well as accredited methods for training with these devices. He is currently the Chief Technology Officer for the Florida Hospital Nicholson Center where he is responsible for establishing the technology strategy and leading technology implementation. He has served as the CTO for the U.S. Army PEO for Simulation, Training and Instrumentation (PEO-STRI); VP and CTO for training systems at Titan Corp; and Vice President of Technology at BTG Inc. He holds a Ph.D. in Computer Science, a Doctorate in Management, and an M.S. in Statistics. He has published 3 professional textbooks on simulation, 10 book chapters, and over 100 journal and conference papers. His most recent book is *Innovation for Innovators: Leadership in a Changing World*. He has served on the editorial boards of the *Transactions on Modeling and Computer Simulation* and the *Research Technology Management* journals.

Comparison of the Usability of Robotic Surgery Simulators

Alyssa Tanaka, M.S., Courtney Graddy, M.S., Roger Smith, Ph.D.
Florida Hospital Nicholson Center
Celebration, FL
Alyssa.tanaka@flhosp.org, courtney.graddy@flhosp.org, roger.smith@flhosp.org

Haider M. Abdul-Muhsin, M.D.
Mayo Clinic
Scottsdale, AZ
Hma980@gmail.com

INTRODUCTION

Robotic surgery has introduced a new dimension into the surgical field. With the introduction of robotic technology between patient and surgeon, a need to master new skills has emerged. Medicine has come to the conclusion that the Halstedian training model (See one, do one, teach one) is no longer sufficient for teaching complex skills, especially robotic surgical skills (Cameron, 1997). A number of simulators have been developed to support training and skill assessment in robotic surgery. The currently available dedicated robotic simulators include: the da Vinci Skills Simulator (dVSS) by Intuitive Surgical Inc., also known as the "Backpack Simulator"; the dV-Trainer from Mimic Technologies Inc.; and the RoSS by Simulated Surgical Sciences LLC (Figure 1). The purpose of these simulators is to train surgeons prior to using the actual system and to allow them to acquire the necessary robotic skills to perform a safe surgery. All of these da Vinci simulators utilize a visual scene that is presented in a computer generated 3D environment providing challenging tests for practicing dexterity and machine operations. Originally, the simulated exercises trained basic robotic skills; however with advances in technology, surgeons can now train for specific procedures (e.g. nephrectomy and hysterectomy).

Figure 1. Simulators of the da Vinci robotic surgical system

Our hospital research laboratory has purchased each of these three simulators for the purpose of studying their effectiveness and applying them to the education of robotic surgeons, specifically for the Department of Defense (DoD). The DoD is interested in the effectiveness of the simulators to train military surgeons prior to and after returning home from deployments. This research is structured as three distinct stages.

From the first stage of this work, the authors summarized the objective characteristics of the three systems. This included descriptions of the exercises offered in each, metrics used to evaluate students, overview of the system administration functions, physical dimensions and configurations of the equipment, and comparisons of the costs of the devices and their support equipment (Smith & Truong, 2013). In the first simulator, the trainee sits at and operates the simulated environment using the actual da Vinci surgical console. The simulator is a custom computer appended to the surgical console through the actual surgical data port. While the simulator costs approximately

$100,000, the surgical console costs $500,000 incurring an investment of $600,000. Using this simulator, users can train using the actual hardware they would use during surgery; however, this requires the use of the surgical console that may be needed to conduct surgeries. Most hospitals may not have a dedicated training console, meaning that users would not have appropriate access to the simulator. The second is a standalone system that utilizes a graphic/gaming computer, connected to a custom desktop viewing and control device that replicates the hardware of the da Vinci surgeon's console. This system shares similar software with the dVSS, but does not require the use of any actual da Vinci hardware. The cost of this simulator is approximately $100,000. The third is composed of a completely customized replica of the da Vinci surgeon's console. Internally the simulator contains a graphic computer, a 3D monitor, and commercial Omni Phantom haptic controllers. This simulator uses unique software and is a little more than $100,000 (Smith &Truong, 2013).

This paper reports on the second stage of this research, in which the validity and usability of the simulators is examined. The third stage will be a measure of learning effectiveness using the systems.

Validity in Surgical Simulation

The validity of medical and surgical simulators is usually measured by the categories defined by McDougal (2007). This paper defines the most commonly recognized forms of validation as: *face, content, construct, concurrent, and predictive validity*. *Face validity* is typically assessed informally by users and is used to determine whether the simulator is an accurate representation of the actual system (i.e. the realism of the simulator). *Content validity* is the measure of the appropriateness of the system as a teaching modality. Experts who are knowledgeable about the device typically assess this via a formal evaluation. *Construct validity* is the ability of a simulator to differentiate between the performances of experienced users and those who are novices. *Concurrent validity* is the extent to which the simulator correlates with the "gold standard" and *predictive validity* is the extent to which the simulator can predict a user's future performance. Collectively, concurrent and predictive validity are known as criterion validity and are used as measures of the simulator's ability to correlate trainee performance with their real life performance. Face and content validity are most effective in evaluating the ability of a simulator to train a surgeon; however construct, concurrent, and predictive validity are most useful for evaluating the effectiveness of a simulator to assess a trainee.

The validity of all three simulators has been tested and reported separately for the da Vinci skill simulator (Hung, Zehnder, Patil, 2011; Kelly, Margules, Kundavaram, 2012; Liss, Abdelshehid, Quach, 2012), the dV-Trainer (Kenney, Wszolek, Gould, Libertino, Moinzadeh, 2009; Sethi, Peine, Mohammadi, 2009; Lee, Mucksavage, Kerbl, 2012) and the RoSS (Seixas-Mikelus, Kesavadas, Srimathveeravalli, 2010; Stegemann et al., 2013; Colaco, Balica, Su, 2012; Raza et al., 2013). To our knowledge only one publication has compared features of two of the simulators, but no comparative studies have been performed with all three of the systems (Liss MA, Abdelshehid C, Quach S., 2012). Thus, the current study aimed to compare all three commercially available da Vinci simulators and detail the findings for face, content, and construct validity for the three systems.

METHODS

Recruitment

Participants in this study included medical students, residents, fellows, and attending physicians. Participants were recruited from the University of Central Florida Medical School, courses held at the Nicholson Center, and two medical robotic conferences (World Robotics Gynecology Congress and Society of Robotic Surgeons Scientific Meeting). Subjects were excluded from participating if they indicated that they had participated in a formal robotic simulation-training course.

Each participant was categorized into one of three groups (i.e. Expert, Intermediate, or Novice) according to the self-reported number of robotic cases (i.e. procedures) he or she had performed. Individuals performing 0-19 robotic cases in which they had 50% or greater console time were categorized as Novices, individuals with 20-99 robotic cases were considered to be Intermediates, and individuals with 100 or more cases were considered to be Experts.

Materials

After being categorized into an experience level, each participant was assigned a specific order in which they used each of the simulators (Figure 2). This order system was used to identify and potentially eliminate any bias that may exist by using a specific system first. All participants completed one exercise on each of the simulators. The tasks chosen were Peg Board 1 in both the dV-Trainer and the dVSS and Ball Placement 1 in the RoSS. The same task was used for both the dV-Trainer and the dVSS because these systems share similar software and exercises. The RoSS software contains unique exercises and Ball Placement 1 is designed to teach the same skills as Peg Board 1.

Participant	Simulator Order
01	CBA
02	BAC
03	ACB
04	CBA
05	BAC
06	ACB
07	CBA
08	BAC
09	ACB

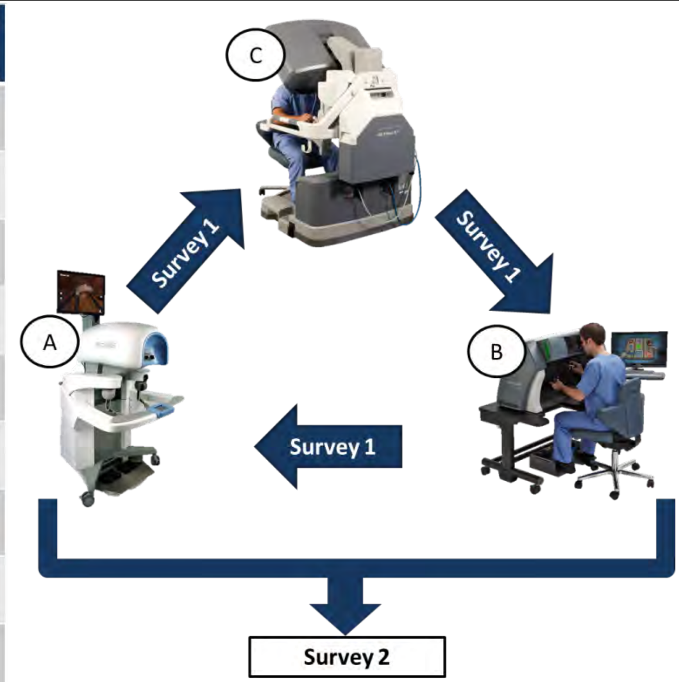


Figure 2. Rotating order of use by subjects, with survey order.

After each exercise on each simulator, participants completed a post questionnaire (Survey 1), which asked for feedback regarding their experience on that specific simulator. After using all three systems, subjects completed a second post questionnaire (Survey 2), which asked them to compare all three systems to each other. The participant's performance metrics were also collected from each of the simulators.

RESULTS

Demographics

Subjects were categorized as Novice (n=37), Intermediate (n=31), or Expert (n=37). Sixty-two percent of subjects were men and 38% were women with an average age of 43. On average, participants had 15 years in practice and 3 years of robotic experience. Seventy-six percent were attending physicians and 73% of participants were currently or had received robotic training, while 41% provided that they train residents and fellows. There were differences in the average age and number of years in practice of participants based on the classification of expert, intermediate or novice (number of robotic procedures). These are to be expected, since higher ages are required to achieve higher number of years of practice and larger numbers of robotic procedures.

Validation

The types of validity evaluated in this experiment were face, content, and construct. To analyze the systems for face validity and content validity, questions from Survey 1 were used. The questions were evaluated on a five point Likert scale (Strongly Disagree, Disagree, Neither Agree or Disagree, Agree, and Strongly Agree). Face validity was analyzed by expert and intermediate feedback as recommended by Van Nortwick et al. (2010) because these are the users most familiar with the robotic system; however, only expert feedback was used for content validity because they have the best ability to judge the appropriateness of the system as a training tool. For construct validity, performance metrics such as Overall Score, Time to Complete, Number of Errors, and Economy of Motion were analyzed (Table 1).

Table 1. Questions and data used for different levels of validity.

Type of Validity	Evaluation	Type of Participant	Question/Metric
Face Validity	Survey 1	Expert and Intermediate	Q1: The hand controllers on this simulator are effective for working in the simulated environment (Likert).
			Q4: The device is a sufficiently accurate representation of the real robotic system (Likert).
Content Validity	Survey 1	Expert	Q2: The 3D graphical exercises in the simulator are effective for teaching robotic skills (Likert).
			Q5: The scoring system effectively communicates my performance on the exercise (Likert).
			Q6: The scoring system effectively guides me to improve performance on the simulator (Likert).
Construct Validity	Simulator	Experts and Novices	Overall Score (points)
			Number of Errors (count)
			Time to Complete (seconds)
			Economy of Motion (centimeters)

Face Validity

The responses of Intermediate and Expert participants (n=68) were used to determine face validity (Table 2). A Chi-square test of independence was used to evaluate the distribution of scores for a specific simulator in relation to the order of the system's presentation to the subject. This analysis indicated that there was no difference in participants' answers according to the order in which the systems were presented; and established that no bias was present due to the presentation order ($p>0.05$). These questions asked participants to evaluate whether the hand controllers on the simulator were effective for working in the simulated environment (Question 1) and if the device is a sufficiently accurate representation of the real robotic system (Question 4). For both questions, the RoSS had the lowest average score, dV-Trainer had the second highest score, and the dVSS had the highest score of the three. A repeated measures ANOVA verified that the systems were scored differently for both questions ($p<0.001$).

Table 2. Average scores from a 5-point Likert scale on face validity.

	DVSS	dV-Trainer	RoSS
Q1: The hand controllers on this simulator are effective for working in the simulated environment.	4.80	3.62	2.17
Q4: The device is a sufficiently accurate representation of the real robotic system.	4.65	3.45	1.82

Content Validity

Expert (n=34) responses were used to determine whether the simulators were appropriate teaching modalities (Table 3). As seen in Table 3, 100% of participants either agreed or strongly agreed that the 3D graphical exercises in the dVSS were effective for teaching robotic skills while 59% disagreed or strongly disagreed that the RoSS' capabilities were effective. When asked if the scoring system effectively communicated their performance, 88% of dVSS users agreed or strongly agreed, while 79% of dV-Trainer users agreed or strongly agreed. Similarly, 91% and

82% of participants agreed or strongly agreed that the dVSS and dV-Trainer, respectively, effectively guided them to improve their performance, while only 36% felt the RoSS provided the same guidance.

Table 3. Scores on a 5 point Likert scale for content validity questions.

Likert Score	**Strong Dis**	**Disagree**	**Neither**	**Agree**	**Strong Agree**
Q2: *The 3D graphical exercises in the simulator are effective for teaching robotic skills.*					
DVSS	0%	0%	0%	35.3%	64.7%
dV-Trainer	2.9%	5.9%	11.8%	50.0%	29.4%
RoSS	20.6%	38.2%	17.6%	17.6%	5.9%
Q5: *The scoring system effectively communicates my performance on the exercise.*					
DVSS	2.9%	5.9%	2.9%	38.2%	50.0%
dV-Trainer	2.9%	2.9%	14.7%	55.9%	23.5%
RoSS	17.6%	20.6%	26.5%	29.4%	5.9%
Q6: *The scoring system effectively guides me to improve performance on the simulator.*					
DVSS	0%	0%	8.8%	61.8%	29.4%
dV-Trainer	2.9%	2.9%	11.8%	61.8%	20.6%
RoSS	18.2%	18.2%	27.3%	33.3%	3.0%

Construct Validity

The overall score, number of errors, time to complete, and economy of motion scores collected by the simulators for Experts (n=37) and Novices (n=37) were used to compare construct validity (Table 4). Overall score is a metric synthesized by multiple metrics and is specific to the individual simulator. Intermediate subjects were not included in the construct validity analysis because it was only necessary to look if the simulator could distinguish specifically between novice and expert users.

For the RoSS, the analysis has 23 missing data points because the system does not report scores when a user exceeds a maximum exercise time or chooses to terminate the exercise before completion. This resulted in a sample of 30 experts and 21 novices on that system. A Mann-Whitney U test showed that the distributions of time ($p=0.221$), number of errors ($p=0.644$), and economy of motion ($p=0.566$) were not statistically different for the experts compared to the novice group. The overall score metric is not automatically exported by the simulator and therefore was not analyzed for this sytem.

The dV-Trainer analysis of experts (n=37) and novices (n=37) had three missing values for economy of motion and completion time and five for the overall score metric, thus the analysis contained varying number of subjects. A Mann-Whitney U test showed that the distribution of the overall scores was not significantly different for the expert compared to the novice group ($p=0.061$). These tests did confirm statistical differences for economy of motion ($p<0.001$) and time to complete ($p<0.001$) for this system with a lower economy of motion value and shorter completion time for expert users compared to novices.

The dVSS analysis included all novice (n=37) and expert (n=37) participants. Using a Mann-Whitney U test, time to complete ($p<0.001$) and overall score ($p=0.006$) were significantly different for the expert compared to the novice group. The expert group had a higher score and a shorter completion time compared to the novice group. However, economy of motion did not show a statistical difference with this analysis ($p=0.216$).

Table 4. Mann-Whitney U test level of significance on construct validity measures

	DVSS	dV-Trainer	RoSS
Time to Complete	p<0.001	p<0.001	p=0.221
Overall Score	p<0.01	p=0.061	n/a
Economy of Motion	p=0.216	p<0.001	p=0.566
Number of Errors	n/a	n/a	p=0.644

The construct validity of the simulators was more specifically analyzed in terms of the self-reported number of cases of all participants (n=105) using a non-parametric correlation coefficient (Spearman's). For the RoSS, 30 participants were excluded from the analysis. For the participants that were included in the analysis (n=75), there was not a significant correlation between time to complete (p=0.181), number of errors (p=0.563), or economy of motion (p=0.390) with the total number of robotic cases performed.

For the dV-Trainer, four participants were excluded from the entire analysis and two participants were excluded from the overall score (Overall Score n=99; Economy of Motion and Time to Complete n=101). When analyzing the number of participants' robotic cases, there was a statistically significant correlation between overall score (p=0.03), economy of motion (p<0.01), and time to complete (p<0.01). The correlation value was negative for economy of motion and time to complete, showing that with a greater number of robotic cases, the time taken and distance moved decreased. The correlation was positive for overall score indicating that the participants' score increased with the number of robotic cases performed.

For the dVSS, two participants were excluded from the analysis (n=103). When analyzing the metrics in terms of the total number of robotic cases performed, there was a statistically significant difference between overall score (p =0.01) and time to complete (p <0.01). The correlation value was negative for time and positive for overall score, signifying that with more robotic cases the time taken decreased and the score increased. There was not a statistically significant correlation between economy of motion and the total number of robotic cases performed (p=0.105).

Table 5. Correlation between level of experience and simulator scores

	DVSS	dV-Trainer	RoSS
Overall Score	p=0.001	p=0.031	n/a
Time to Complete	p<0.001	p<0.001	p=0.181
Economy of Motion	p=0.105	p<0.001	p=0.390
Number of Errors	n/a	n/a	p=0.563

Usability (Preference)

The questions from the Survey 2 were used to understand the preference of the subjects when using the simulators. All subjects were included in this analysis except for two participants who were dropped from the analysis because they did not complete the questionnaire. The participant's responses to the usability questions can be seen in Figure 3:

- *If you are (were) a program director, which simulator would you choose for your trainees;*
- *In which simulator were you physically more comfortable;*
- *Which simulator had the best hand controls;*
- *Which simulator had the best foot controls;*
- *Which simulator had the best 3D vision;*
- *Were you feeling stressed or annoyed by any of the simulators?*

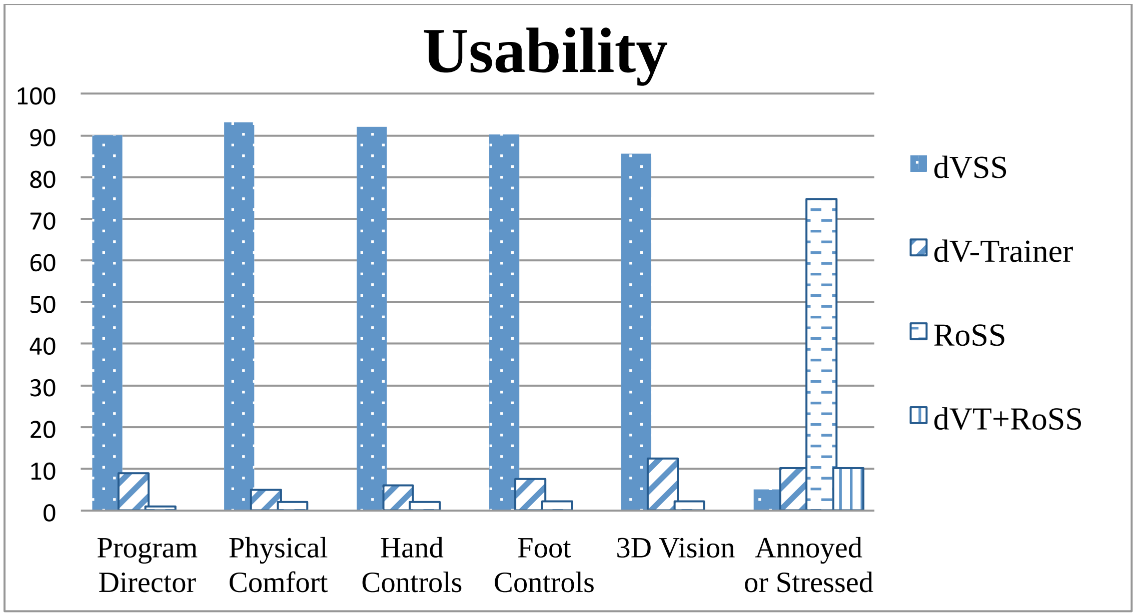


Figure 3. Description of usability responses

Overall, most participants preferred the dVSS and indicated that they would choose this device as a training system if they were a program director. Participants not only felt most comfortable in the dVSS, but also felt that the system had the best control and vision equipment. The least preferred system was the RoSS which most participants also agreed made them feel stressed or annoyed. Ten percent of participants also responded that they felt stressed or annoyed by both the dV-Trainer (dVT) and the RoSS.

Cost

All participants were also asked to provide feedback on their simulator preference in terms of the cost of the system. The responses were analyzed in terms of the frequency of the responses given. Most participants felt that the mimic dV-Trainer was worth the investment; while most felt that the RoSS was not worth the money. When asked about the dVSS, only 56% of participants agreed that it was worth the investment. Figure 4 provides a full description of the responses.

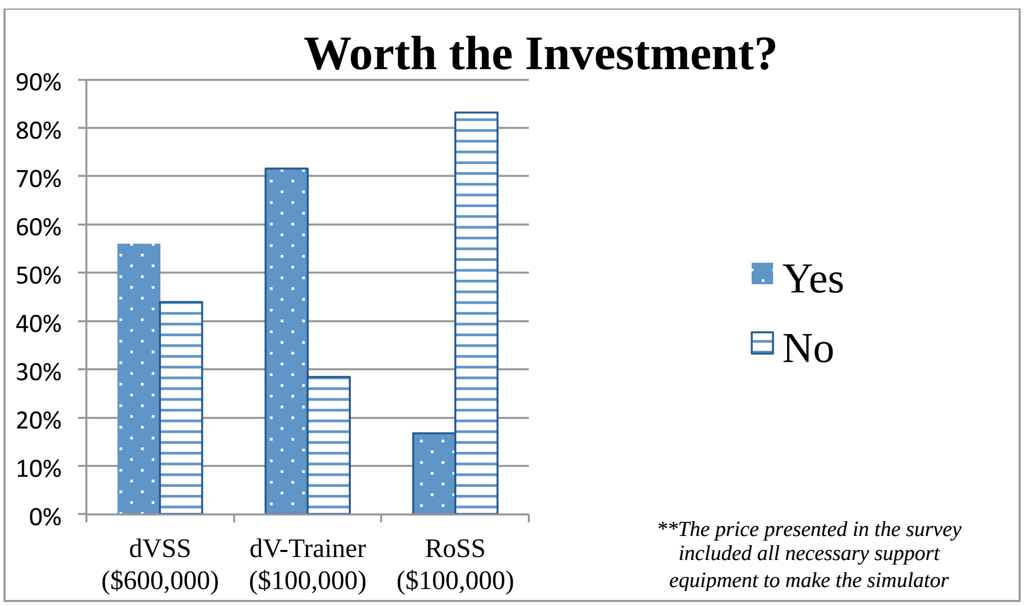


Figure 4. Description of cost preferences

DISCUSSION

The aim of this study was to conduct a comparison of the three commercially available simulators used to train surgeons on the daVinci robotic system. The study was performed for the US Army to assist them in making a purchasing and deployment decision regarding robotic simulators. Their interest is in re-training robotic surgeons who have been deployed to combat zones, where they have served as trauma surgeons for many months. Prior to resuming their robotic specialties, these surgeons need a program to both refresh and re-validate their robotic skills. This study provided information about the face, content, and construct validity as well as usability of the systems. The simulators were perceived to be different in their representation of the real robotic system. The dVSS was most preferred in terms of ergonomics and usability; however, most participants did not feel that this system was worth a $600,000 investment. In terms of cost, most participants agreed that the dV-Trainer had the best cost-effectiveness. The RoSS was the least preferred system for comfort and other usability aspects (i.e., hand controls, foot controls, and 3D interface), with most participants feeling stressed or annoyed when using the system. This study was unable to validate the face, content, or construct validity for this system.

The dVSS leverages the actual hardware used to perform robotic surgeries for use in the simulated environment, which allows for a more realistic experience, but decrease its availability and creates a higher cost for training than other robotic simulators. Economy of motion was not able to differentiate novices from experts in the dVSS, which could be attributed to the ease of use of the controllers allowing novices to move the controls as efficiently as experts. The generous workspace of the dVSS could also have an impact on the lack of difference. In contrast to the dVSS, the dV-Trainer is a standalone simulator and does not require the support of the daVinci hardware to operate. This allows for better accessibility and requires less of an investment for training. The overall score aspect of construct validity may not have shown a difference between novices and experts because of the way that the scoring is developed. The scoring system is constructed with a “ceiling” that prevents users from achieving a high overall score without attaining high scores across multiple metrics.

Currently, there is limited data available that confirms construct validity of the RoSS. Similarly to Raza (2013), this study was unable to confirm a difference between experts and novices in terms of time taken to complete the exercise. Time to complete, as well as economy of motion, is considered a highly relevant measurement of expertise levels for robotic surgeons (Perrenot, Perez, Tran, Jehl, Felblinger, Bresler, & Hubert, 2012). To our knowledge this three-part study is the first to compare all three available systems. This study involved the largest sample size and diversity of participants (i.e., experience levels, number of robotic cases, and subspecialty type) thus far in relevant publications. The lack of consistency in the available exercises and scoring systems across the three systems was a limitation to the study. Considerations for future research would be to use more complex exercises and increase the depth of the face and content validity evaluation.

Current research is focused on the effectiveness of the simulators and objectively measuring the transfer of training to the actual robotic system. All three simulators will be examined in this final stage of the experiment; however, the results of this three-part study will guide the choice of simulators used for future studies at Florida Hospital Nicholson Center and may also influence decisions at other laboratories. Also, this research may impact the purchasing decisions of customers for these devices.

REFERENCES

Cameron, J. L. (1997). William Stewart Halsted. Our surgical heritage. *Annals of surgery*, *225*(5), 445.

Colaco M, Balica A, Su D. (2012) Initial experiences with RoSS surgical simulator in residency training: a validity and model analysis, J Rob Surg, 7(1):71-75

Hung AJ, Patil MB, Zehnder P. (2012). Concurrent and predictive validation of a novel robotic surgery simulator: a prospective, randomized study. J Urology, 187(2):630-7.

Hung AJ, Zehnder P, Patil MB. (2011). Face, content and construct validity of a novel robotic surgery simulator. J Urology, 186(3):1019-24.

Kelly DC, Margules AC, Kundavaram CR. (2012). Face, content, and construct validation of the da Vinci Skills Simulator. J Urology, 79(5):1068-72.

Kenney P, Wszolek M, Gould J, Libertino J, Moinzadeh A. (2009). Face, Content, and Construct Validity of dV-Trainer, a Novel Virtual Reality Simulator for Robotic Surgery. J Urology, 73(6):1288-92.

Lee JY, Mucksavage P, Kerbl DC. (2012) Validation study of a virtual reality robotic simulator--role as an assessment tool? J Urology, 187(3):998-1002.

Liss MA, Abdelshehid C, Quach S. (2012). Validation, Correlation, and Comparison of the da Vinci Trainer(™) and the da Vinci Surgical Skills Simulator(™) Using the Mimic(™) Software for Urologic Robotic Surgical Education. J Endourol, 26(12):1629-34.

McDougall EM. (2007). Validation of surgical simulators. J Endourol, 21(3):244-7.

Sethi AS, Peine WJ, Mohammadi Y. (2009). Validation of a novel virtual reality robotic simulator. J Endourol, 23(3):503-8.

Perrenot, C., Perez, M., Tran, N., Jehl, J. P., Felblinger, J., Bresler, L., & Hubert, J. (2012). The virtual reality simulator dV-Trainer® is a valid assessment tool for robotic surgical skills. *Surgical endoscopy*, *26*(9), 2587-2593.

Raza SJ, Chowriappa A, Stegemann A, Ahmed K, Shi Y, Wilding G, Kaouk J, Peabody J, Menon M, Kesavadas T, Guru KA. (2013). Construct Validation of the Fundamental Skills of Robotic Surgery (FSRS) Curriculum. Podium presentation, AUA Annual Congress, San Diego,

Seixas-Mikelus SA, Kesavadas T, Srimathveeravalli G. (2010). Face validation of a novel robotic surgical simulator. J Urology, 76(2):357-60.

Smith, R., & Truong, M. (2013, January). Robotic Surgical Education with Virtual Simulators. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* (Vol. 2013, No. 1). National Training Systems Association.

Stegemann A, Kesavadas T, Rehman S, Sharif M, Rao M, duPont N, Shi Y, Wilding G, Hassett J, and Guru K. (2013). Development, Implementation, and Validation of a Simulation-Based Curriculum for Robot-Assisted Surgery. J Urology, 81(4):767-74.

Van Nortwick, S. S., Lendvay, T. S., Jensen, A. R., Wright, A. S., Horvath, K. D., & Kim, S. (2010). Methodologies for establishing validity in surgical simulation studies. *Surgery*, *147*(5), 622-630.

Poster Title:

Fundamentals of Robotic Surgery Psychomotor Skills: Metrics Development and Evaluation

Introduction:

Robotic surgery has been established as an innovative approach in surgery due to a telemanipulator device, which introduced a new dimension into surgical tools. This device allows surgeons to manipulate robotic arms from a remote console to perform complex surgical procedures. Robotic surgical systems overcome laparoscopic limitations and facilitate the performance of minimally invasive surgery due to 3D vision, 7-degree-of-freedom instruments, tremor abolition, motion amplification, and stabilization of the camera (Patel et al., 2013; Hubens, Coveliers, Balliu, Ruppert, & Vaneerdeweg, 2003; Blavier, Gaudissart, Cadière, & Nyssen, 2007). The system also offers 10x magnification, wristed instruments, and a third working arm. Currently, the only system is Intuitive's da Vinci Surgical System.

Robotic surgery has demonstrated safety and effectiveness for urologic, gynecologic, ENT, and complex general surgery procedures (Barbash, Friedman, Glied, & Steiner, 2014; Serati et al., 2014; Maan, Gibbins, Al-Jabri, & D'Souza, 2012; Luca et al., 2013; Zureikat et al., 2013). Exponential growth of minimally invasive procedures, particularly robotic-assisted procedures, raises the question of how to assess robotic surgical skills. This device also introduces a specific need for training and certification to ensure a minimal standard of care for all patients. Some institutions have attempted to develop and validate robotic training in regards to specific specialties (Chitwood et al., 2001; Geller, Schuler, & Boggess, 2011; Grover, Tan, Srivastava, Leung, & Tewari, 2010; Chowriappa et al., 2014; Jarc & Curet, 2014); however, the lack of a national standard has pushed surgical societies (e.g. SAGES, SRS, and MIRA) to develop a unified approach and standard for robotic skills training Zorn et al., 2009).

To develop a comprehensive model for robotic surgery, the Department of Defense, Veterans Administration, and fourteen surgical specialty societies convened multiple consensus conferences to create the Fundamentals of Robotic Surgery (FRS). A similar education and training initiative was implemented for use in laparoscopic surgery, which resulted in a curriculum called Fundamentals of Laparoscopic Surgery (FLS). FRS Conference participants included more than 80 subject matter experts (SMEs), consisting of surgeons, psychologists, engineers, simulation experts, and medical educators (Smith, Patel, Chauhan, & Satava, 2013).

The committee's vision of FRS was driven by two main goals: to ensure a perfect understanding of the basics of robotic surgery and to develop a psychomotor skills program that focused on basic robotic tasks. The intended users for this program are novice robotic surgeons, who could be residents or fellows and attending surgeons who have never used the robotic system. Two assessment tools were created: an online curriculum for knowledge and team training skills and a device for psychomotor skill training and evaluation (Levy, n.d.). For the psychomotor skills portion of the training, physical and virtual devices were developed (Figure 1).

<image of physical and virtual domes>

Objective:

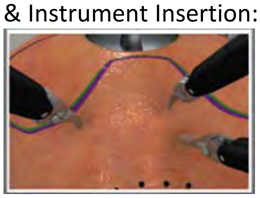
The purpose is to describe the development of the exercise metrics and how they are designed to take into account the specificity of robotic errors.

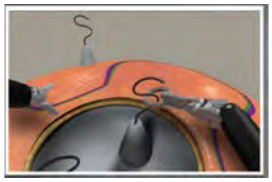
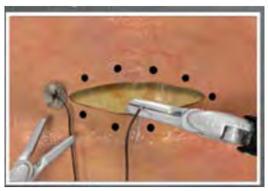
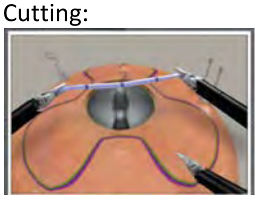
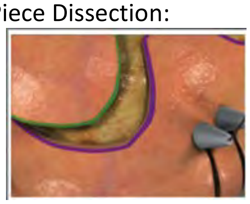
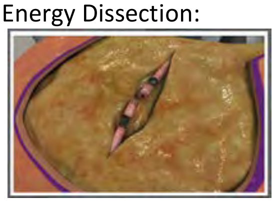
Methods:

The members of the consensus conferences worked together to outline outcomes measures and metrics, which touched on the essential cognitive, psychomotor, and team training skills. This resulted in a prioritized matrix of 25 robotic surgery concepts, which is the core material used in the design and development of the FRS Curriculum. Of those 25 concepts, 16 are directly linked with psychomotor skills (Table 1). The FRS committee members then identified seven exercises that incorporated all 16 skills.

Phase	Items	
Pre-Operative	Docking	Robotic Trocar
Intra-Operative	Energy Sources	Clutching
	Camera Control	Wrist Articulation
	Instrument Exchange	Multi-Arm Control
	Eye-hand coordination	Atraumatic Handling
	Dissection (Fine and Blunt)	Cutting
	Needle Driving	Suture Handling
	Knot Tying	
Post-Operative	Undocking	

These exercises include docking and instrument insertion, tower transfer, knot tying, railroad track, 4th arm cutting, puzzle piece dissection, and vessel energy dissection (Table 2). *Docking and instrument insertion* is an essential and unique robotic skill to begin a procedure. Failure at this stage of the procedure can compromise the surgery. *Tower transfer* is a non-surgical exercise that introduces the utilization of endowrist manipulation and the 7 degrees of freedom to surgeons. *Knot tying* and *railroad track* are the base of a suturing exercise. The technology introduced in the wristed instruments facilitates the performance of these tasks. *4th arm cutting* is another task specific to robotics, which improves surgeon's autonomy. The 4th arm allows surgeons to manage three instruments by using a foot pedal to switch between working arms. *Puzzle piece* and *vessel energy dissection* are critical tasks, which incorporate complex articulation of instruments and application of energy (i.e. cauterization and cutting).

Exercises	Skills
Task1: Docking & Instrument Insertion: 	- Docking - Instrument insertion - Eye-hand coordination - Operative field of view
Task 2: Ring Tower Transfer:	- Eye-hand coordination - Camera navigation - Clutching - Wrist articulation - A-traumatic handling

	
Task 3: Knot Tying: 	- Knot tying - Suture handling - Eye-hand coordination - Wrist articulation
Task 4: Railroad Track: 	- Needle handling & manipulation - Wrist articulation - A-traumatic handling - Eye-hand coordination
Task 5: 4th Arm Cutting: 	- Multiple arm control & switch - Cutting - A-traumatic handling - Eye-hand coordination
Task 6: Puzzle Piece Dissection: 	- Sharp and blunt dissection - Cutting - A-traumatic handling - Eye-hand coordination - Wrist articulation
Task 7: Vessel Energy Dissection: 	- Energy sources use - Sharp dissection - Cutting - Multiple arm control - A-traumatic handling - Eye-hand coordination

After establishing the exercises, metrics were created to assess the users task specific performance. The development of these metrics began early in the device development process. To do this, the team consulted with experts robotic surgeons to determine what would constitute performing each exercise correctly. Often this required looking at the purpose of the exercise and what skills are being taught. By doing this, the team was able to outline metrics that were relevant to the task and important to robotic skills. We then evaluated how certain mistakes, which are specific to the FRS exercises, are translatable to robotic skill. For example, in ring tower transfer and knot tying, excessive instrument force can be evaluated in the form of knocking a tower off of the surface. Other metrics like the time to complete the task, the number of times the instruments are placed out of view, and the number of instrument collisions were also considered in the development.

After the basic metrics were established, the team evaluated an expert robotic surgeon and a novice using the metrics. The expert helped the team to envision how the exercises should be performed, while the novice testers helped to see what mistakes are typically made by inexperienced users. The metrics were also tested using different types of materials with different attributes to ensure that changes in the materials would not limit the user's ability to achieve the desired metrics. For example, it was important that the synthetic fat material have a slightly sticky texture to evaluate the users performance in dissecting the skin layer from the fat.

Once established, it was important to consider how the metrics would be measured. Since the scoring would be performed via expert video review, the research team had to consider the feasibility of the scoring for the reviewer. For example, it was not realistic to ask the reviewer to account for the seconds that the instruments were out of view or attempt to measure the number of centimeters the user cut outside a designated line. So, for these metrics it was decided to use a count or a yes/ no type response. Also, some metrics were added to ensure that the exercise was performed correctly. For example, in 3rd arm cutting exercise, a metric was added to ensure that the user performed the cuts in the correct order. This is relevant because when performed in another order, the user does not actually have to practice switching back and forth between arms.

After the physical dome metrics were established, they were given to simulation companies to use as metrics for evaluating the simulated dome exercises. When working with the simulation teams for development of the virtual domes, it was important to ensure that the virtual domes mimicked the physical properties of the actual dome. This was highly important because the simulations may be used without a proctor and the physical behaviors have a considerable impact on the scoring metrics and guidance that is given for improving performance. The research team worked with simulation developers to evaluate the simulated exercise properties including elasticity of materials, flexibility of sutures, simulated gravity, and the effects of excess force on the virtual device to ensure that it behaved similar to the real dome. Thus, creating similar metrics across all dome modalities. Many "what if" situations were examined to determine how they would be handled in the real dome and how the simulation should handle them. These incidences could have different implications for instruction and assessment in a real dome exercise, as opposed to in the simulated environment. Finally, it was an important consideration to leverage the capabilities of the simulated environment to objectively assess the users performance.

Results:

The physical FRS dome is scored using a validated global assessment called the Global Evaluative Assessment of Robotic Skills (GEARS). This looks at the overall performance of the user in terms of their robotic skills. The described work resulted in task specific metrics for each exercise (Table 3).
<insert metrics table>

Most of the metrics used to evaluate the virtual dome are similar to the metrics of the physical dome. However, some of the metrics do vary between the physical and simulated domes. While the physical dome is scored via expert video reviewing, the simulator can more objectively assess a user's performance. This allows the simulated exercises to score some errors more accurately, such as instruments being out of view for a specific amount of time (seconds) and over a specific distance (centimeters).

Conclusions:

Many considerations were made during the development of the FRS psychomotor skills metrics. From this process, we have taken away several lessons that helped to develop our metrics including:

- Perform preliminary testing with varying expertise level users
- Create specific and measurable standards for assessment
- Work closely with the simulation development teams for the details of simulation
- Constant testing and expert consultation

It was important to determine how users of varying expertise level would use the device, including determining how the expert would perform it and what mistakes are typically made by novices. It was important for us to ensure that the metrics developed were specific and measurable under the circumstances that the users would be assessed. The video reviewing did not allow for certain measures to be assessed causing some metrics for the physical dome to be modified to an easier format for video review. The simulated environments however allow for a much more objective assessment and it was beneficial for our team to work closely with the simulation developers to establish what metrics might change in the virtual modality. Finally, it was very important to solicit expert feedback on the metrics throughout the process.

Further validation testing is necessary for the FRS curriculum. This is currently being assessed via a pilot study and eventually in a multi-site validation study. The goal of the pilot study is to establish preliminary scoring benchmarks based on expert performance, which will be used to guide the multi-site validation study.

References:

Acknowledgements:

The authors thank the members of the FRS committee and design team for the initial concept and design guidance throughout the project as well as development of the FRS curriculum.

CONTROL ID: 2086488

CONTACT (NAME ONLY): Mireille Truong

TITLE: ROBOTIC SURGICAL SIMULATION VERSUS TRADITIONAL DIDACTICS FOR SURGICAL TRAINING: A RANDOMIZED CONTROLLED TRIAL

AUTHORS (FIRST NAME INITIAL LAST NAME): M. D. Truong[1, 2], A. Tanaka[2], C. Graddy[2], K. Simpson[1, 2], M. Perez[2], A. P. Advincula[1, 2], R. D. Smith[2]

INSTITUTIONS (ALL):

1. Obstetrics and Gynecology, Columbia University Medical Center, New York, NY, United States.
2. Research Department, Nicholson Center, Celebration Health at Florida Hospital, Celebration, FL, United States.

ABSTRACT BODY:

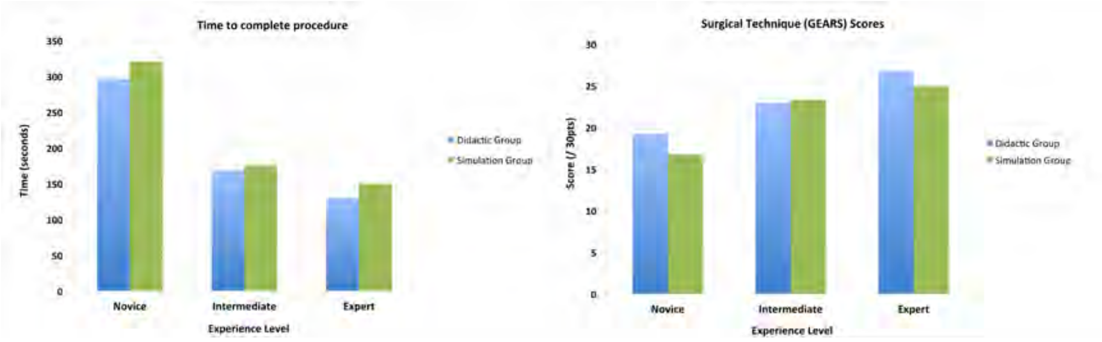
Objectives : To compare the impact of two surgical training methods on surgical performance: traditional didactic lectures versus robotic surgical simulation.

Materials and Methods: Prospective randomized controlled trial conducted at a surgical and medical training center. Medical students, residents, fellows and attendings were recruited from various surgical specialties. Based on surgical experience level, subjects were randomly assigned to either the Simulation Group (SG), which trained on the dv-Trainer® robotic simulator; or to the Didactic Group (DG), which trained via a structured didactic lecture in preparation to perform a robotic porcine cystotomy repair. Each group received the same training duration. Both groups completed a post-training knowledge test and a self-evaluation. Blinded surgeons evaluated the subjects' recorded surgical performances based on ability to complete required task, number of errors and surgical technique using a validated robotic surgical assessment tool (GEARS). Time and performances were compared between the two groups. Data regarding demographics, feedback on both training methods, and preparedness to perform the robotic surgical procedure were collected. Chi-square test and ANOVA were used for statistical analysis. The study was powered at 80% to detect a difference of 15% in performance between the two groups with a p-value of <0.05.

Results : Total n= 125 (52 novices, 42 intermediates, 27 experts). Both groups (DG n=64; SG n=61) were similar in distribution of age, gender and training and experience level. The majority of subjects had prior robotic simulator use. Mean cystotomy repair time was similar in both groups (DG=224min; SG=219min, p=0.138). The overall performances between SG and DG were not significantly different but when controlled for experience level, DG novices performed better than SG as far as number of errors (p= 0.03) while SG experts outperformed DG experts in terms of number of errors (p=0.029). For the post-training knowledge test, the DG scored higher than the SG (DG 84%, SG 72%, $p<0.001$). Both groups felt that neither simulation nor didactic lectures alone would be sufficient for robotic surgical training but that if given a choice between the two, they would choose simulation for their training both before (DG 94%, SG 93%) and after (DG 90%, SG 90%) their respective interventions. Additionally, both groups felt that simulation would be an effective tool for robotic surgical training (DG 88%, SG 77%).

Conclusion : In this study, simulation and didactics training seem to offer similar acquisition of surgical skills, challenging the notion that hands-on surgical simulation is superior to traditional didactic lectures. The results suggest that a well-designed didactic program may be just as effective as simulation and better for cognitive training. Effectiveness of training methods may differ based on experience levels, where simulation may be more effective for experts while didactic lectures may be more effective for novices. Despite the similarity in their performances, both groups preferred simulation to didactics. The effectiveness of combining both methods is yet to be determined.

(no table selected)



Didactic vs Simulation: surgical performances based on level of experience

Presentation Details

PRESENTATION TYPE: Oral - Full or Oral Poster
CURRENT CATEGORY: Surgical Education Technical
Other Abstract Details

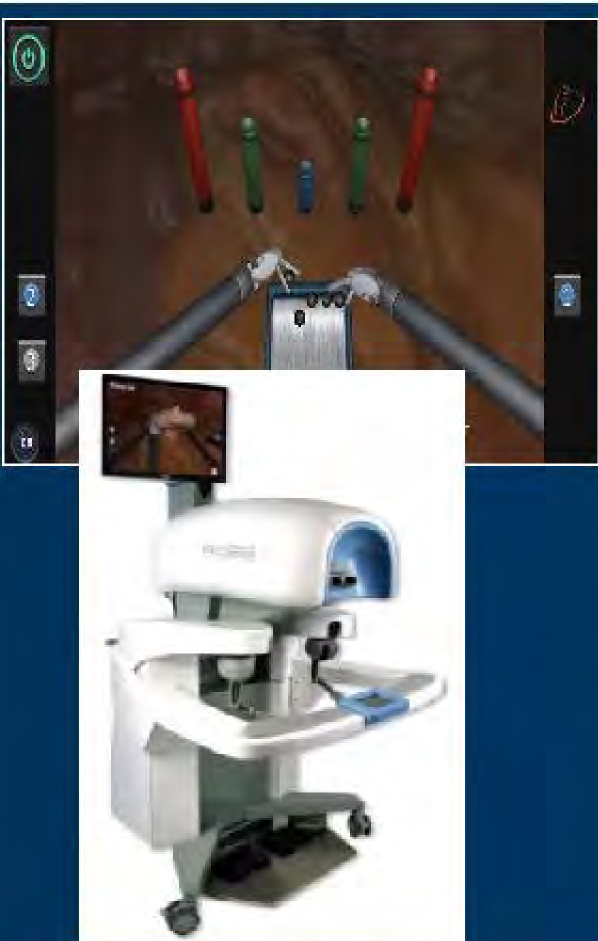
Update on Robotic Surgical Simulation

Roger Smith *PhD*
Florida Hospital Nicholson Center
roger.smith@flhosp.org

http://www.nicholsoncenter.com/

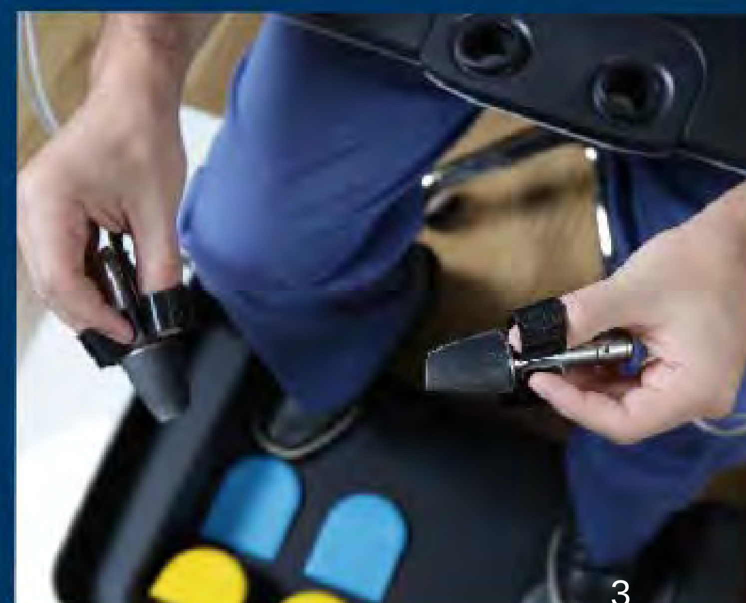
FLORIDA HOSPITAL
NICHOLSON CENTER

Simulator Comparison



Part I: Functionality – feature comparison
Part II: Usability – instructor & student opinions
Part III: Effectiveness – student improvement

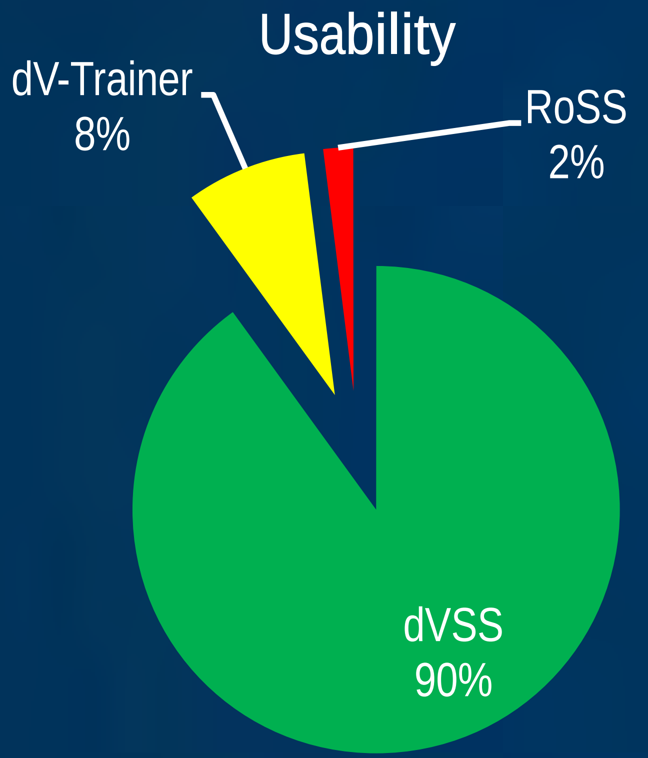
Robotix Mentor (3D Systems/Simbionix)

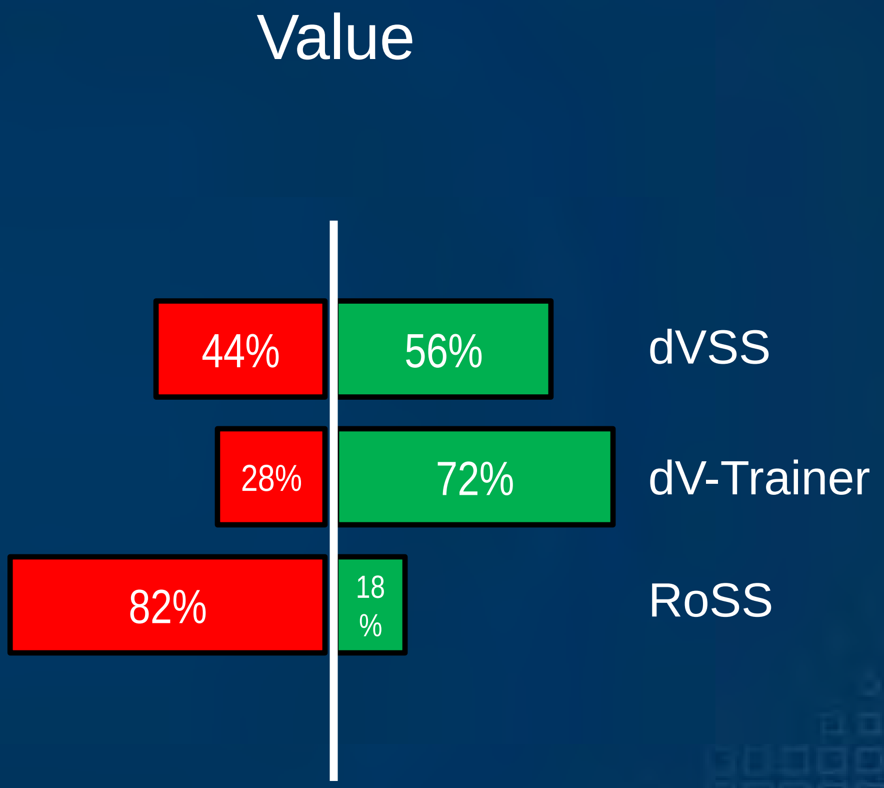


Part I: System Functionality Comparison

Features	DVSS	dV-Trainer	RoSS
System Manufacturer	Intuitive Surgical Inc.	Mimic Technologies Inc.	Simulated Surgical Systems LLC
Specifications (Simulator only)	Depth 7” Height 25” Width 23” 120 or 240V power	Depth 36” Height 26” Width 44” 120 or 240V power	Depth 44” Height 77” Width 45” 120 or 240V power
Specifications (Complete System as shown in Figure 1)	Depth 41” Height 65” Width 40” 120 or 240V power	Depth 36” Height 59” Width 54” 120 or 240V power	Depth 44” Height 77” Width 45” 120 or 240V power
Visual Resolution	VGA 1024 x 768	VGA 1024 x 768	VGA 640 x 480
Components	Customized computer attached to da Vinci surgical console	Standard computer, visual system with hand controls, foot pedals.	Single integrated custom simulation device
Support Equipment	da Vinci Si surgical console, custom data cable	Adjustable table, touch screen monitor, keyboard, mouse, protective cover, custom shipping container	USB adapter, keyboard, mouse
Exercises	40 simulation exercises (35 by Mimic, 5 by Simbionix)	65 simulation exercises	52 simulation exercises.
Optional Software	PC-based Simulation management	Mshare curriculum sharing web site	Video and Haptics-based Procedure Exercises (HoST)
Scoring Method	Scaled 0-100% with passing thresholds in multiple skill areas	Proficiency-based point system with passing thresholds in multiple skill areas	Point system with passing thresholds in multiple skill areas
Student Data Management	Custom control application for external PC. Export via USB memory stick.	Export student data to delimited data file and graphical reports.	Export student data to delimited data file.
Curriculum Customization	None	Select any combination of exercises. Set passing thresholds and conditions.	Select specifically grouped exercises. Set passing thresholds.
Administrator Functions	Create student accounts on external PC. Import via USB memory stick.	Create student accounts. Customize curriculum.	Create student accounts. Customize curriculum.
System Setup	None.	Calibrate controls.	Calibrate controls.
System Security	Student account ID and password.	PC password, Administrator password, Student account ID and password.	PC password, Administrator password, Student account ID and password.
Simulator Base Price	$85,000	$99,200	$126,000
Support Equipment Price	$500,000	$9,800	$0
Total Functional Price	$585,000	$109,000	$126,000

Part II: Usability & Value for Teaching





Robotic Skills vs. Simulator Exercises/Metrics

Core Skills	Content	Simulators		
		dV-Trainer	RoSS	DVSS
Wristed Bimanual Manipulation	Tasks	Yes	Yes	Yes
	Metrics	No	No	No
Maneuver Camera	Tasks	Yes	Yes	Yes
	Metrics	No	No	No
Master Cntrl Clutching	Tasks	No	Yes	Yes
	Metrics	M Wrk Rng	No	M Wrk Rng
Use 3rd Arm	Tasks	Yes	Yes	Yes
	Metrics	No	No	No
Depth/3D Perception	Tasks	No	No	No
	Metrics	No	No	No
Aware Instrument Force	Tasks	No	No	No
	Metrics	Ex Inst Frc	No	Ex Inst Frc

Liu & Curet, *Teaching and Learning in Medicine*, Jan 2015

Robotic Skills vs. Simulator Exercises/Metrics

Core Skills	Content	Simulators: dV-Trainer	RoSS	DVSS	Robotix Mentor	FRS Dome
Wristed Bimanual Manipulation	Tasks	Yes	Yes	Yes	Yes	Yes
	Metrics	No	No	No	No	Yes
Maneuver Camera	Tasks	Yes	Yes	Yes	Yes	Yes
	Metrics	No	No	No	No	Yes
Master Cntrl Clutching	Tasks	No	Yes	Yes	Yes	Yes
	Metrics	M Wrk Rng	No	M Wrk Rng	Yes	No
Use 3rd Arm	Tasks	Yes	Yes	Yes	Yes	Yes
	Metrics	No	No	No	Yes	Yes
Depth/3D Perception	Tasks	No	No	No	No	Yes
	Metrics	No	No	No	No	Yes
Aware Instrument Force	Tasks	No	No	No	Yes	Yes
	Metrics	Ex Inst Frc	No	Ex Inst Frc	Yes	Yes

Robotic Simulator ROI Model

Simulator Investment Costs

- Cost of Simulators
- Annual Maintenance Fee
- Facility Costs
- Staffing Costs
- Supplies & Materials

Supplies & Materials	$1,000	$200	
Investment	$217,000	$60,200	$277,200

Surgeon Productivity

Number of Surgeons Trained	3	
Mean Length of Operation	2	hours
Time Saved per Procedure	10	minutes
Length of Surgical Day	8	hours
Surgeries per day	12	
New Free Surgery Time per day	2.00	hours
Added Surgeries per Day	1	
Revenue per Surgery	$10,000	
Additional Revenue per day	$10,000	
Surgical Days	100	per year
Additional Revenue per Year	$1,000,000	$1,000,000

Surgeon Productivity

- Reduced Surgery Time
- Increased Surgeries per Day
- Surgeon Health
- Increased Surgeon Stamina (hrs/day)

Hospital Costs

- Reduced Liability Insurance
- Reduced Instrument Breakage
- Reduced Surgical Error Rate

Other Training Costs

- Reduced Training in the OR
- Reduced In-house Instructors
- Reduced External Training w Travel

External Training Event Cost (w/ Travel)	$10,000	
Number of External Events	6	per year
Current Training Cost	$210,000	
Training Reduction	50%	
Training Cost Savings	$105,000	$105,000

Robotic Simulator ROI Model

Return On Investment	**Year 1 Investment**	**Year 1 Return**	**Year 1 ROI**	**Year 1 Pay Back**	
	$277,200	$1,128,150	407%	90	days
	Year 2 Investment	**Year 2 Return**	**Year 2 ROI**	**Year 2 Pay Back**	
	$60,200	$1,128,150	1874%	19	days

Fundamentals of Robotic Surgery

Roger Smith *PhD*

Florida Hospital Nicholson Center

roger.smith@flhosp.org

http://www.nicholsoncenter.com/

FLORIDA HOSPITAL
NICHOLSON CENTER

Fundamentals of Robotic Surgery

- The Fundamentals of Robotic (FRS) is a joint educational program funded through a Department of Defense grant and an unrestricted educational grant from Intuitive Surgical.
- FRS is a multi-specialty, proficiency-based curriculum of basic technical skills to train and assess surgeons to safely and efficiently perform robotic-assisted surgery.
- It was developed by over 80 national/international robotic surgery experts, behavioral psychologists, medical educators, statisticians and psychometricians.
- The clinical robotic surgery subject matter experts represented all of the major surgical specialties in the United States that currently performs robotic-assisted surgical procedures, the Department of Defense and the Veterans Administration (VA).

Leadership

Richard Satava, MD
Professor of Surgery
University of Washington
Seattle, WA

Roger D. Smith, PhD
Chief Technology Officer
Florida Hospital Nicholson Center
Celebration, FL

Vipul R. Patel, MD
Medical Director,
Global Robotics Institute
Florida Hospital Celebration Health
Celebration, FL

Robert M. Sweet, MD
Associate Professor Urology
Director of Medical School
Simulation Programs,
University of Minnesota
Minneapolis, MN

Jeffrey S. Levy, MD
CEO, CaseNetwork
Newtown Square, PA

Dimitrios Stefanidis, MD, PhD
Medical Director,
Carolinas Simulation Center
Associate Professor of Surgery,
Carolinas Healthcare System
Charlotte, NC

Arnold Advincula, MD
Professor of Obstetrics & Gynecology
Columbia University,
College of Physicians and Surgeons
New York, NY

Martin A. Martino, MD
Medical Director,
Minimally Invasive Robotic Surgery
Lehigh Valley Health Network
Allentown, PA

Organizations Represented

- American Academy of Orthopedic Surgeons
- American Academy of Otolaryngology – Head and Neck Surgery
- American Association of Gynecologic Laparoscopists
- American Association for Thoracic Surgery
- American Board of Surgeons
- American Board of Urology
- American College of Surgery
- American College of Obstetricians and Gynecologists
- American Hernia Society
- American Society of Colon & Rectal Surgeons
- American Society of Plastic Surgeons
- American Urological Association
- American Urogynecologic Society
- Arthroscopy Association of North America
- Asia Pacific Hernia Society
- Association for Surgical Education
- China Hernia Society
- Latin American Association of Laparoscopic Surgery
- Latin American Hernia Foundation
- Minimally Invasive Robotic Association
- National Institute for Health Research, UK
- Resident Review Committee – Surgery
- Resident Review Committee – Urology
- Royal Australasian College of Surgeons
- Royal College of Surgeons, Ireland
- Society of American Gastrointestinal and Endoscopic Surgeons
- Society of Laparoendoscopic Surgeons
- United States Department of Defense
- Veterans Health Administration, National SimLEARN Center

FRS Participants (partial)

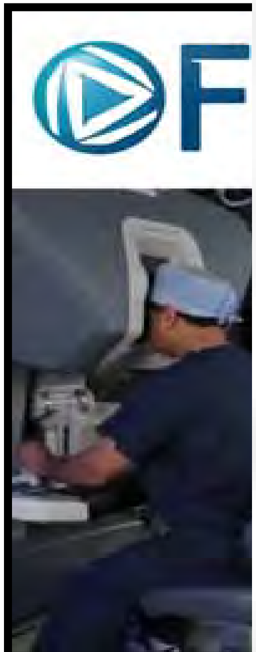
Arnold Advincula, MD, FACS
Rajesh Aggarwal, MBBS
Abdulla Ali Al Ansari, MD, FRCS
David M. Albala, MD
Richard L. Angelo, MD
Mehran Anvari, MD
John Armstrong, MD, FACS
Garth Ballantyne, MD, MBA
Michele Billia, MD
James F. Borin, MD
David M. Bouchier-Hayes, MD
Timothy C. Brand, MD, FACS
Jan Cannon-Bowers, PhD
Sanket Chauhan, MD
Rafael F. Coelho, MD
Geoff Coughlin, MD
Alfred Cuschieri, MD
Prokar Dasgupta, MD
Ellen Deutsch, MD
Gerard Doherty, MD
Brian J. Dunkin, MD, FACS
Susan G. Dunlow, MD
Gary Dunnington, MD
Ricardo Estape, MD
Peter Fabri, MD
Vicenzo Ficarra, MD
Marvin Fried, MD
Gerald Fried, MD
Vicenzo Ficarra, MD
Anthony G. Gallagher, PhD
Larry R. Glazerman, MD, MBA

Teodor Grantcharov, MD, PhD, FACS
Piero Giulianotti, MD
David Hananel
James C. Hebert, MD, FACS
Robert Holloway, MD
Santiago Horgan, MD
Jacques Hubert, MD
Wallace Judd, PhD
Lenworth Jacobs, MD
Arby Kahn, MD
Keith Kim, MD, FACS
Sara Kim, PhD
Michael Koch, MD, FACS
Timothy Kowalewski, PhD
Rajesh Kumar, PhD
Kevin Kunkler, MD
Gyunsung Lee, PhD
Thomas S. Lendvay, MD
Raymond J. Leveillee, MD
Jeffrey S. Levy, MD
C.Y. Liu, MD
Fred Loffer, MD
Guy Maddern, FRACS
Scott Magnuson, MD
Javier Magrina, MD
Michael Marohn, MD
David Maron, MD
Martin A. Martino, MD, FACOG
W. Scott Melvin, MD
Francesco Montorsi, MD
Alex Mottrie, MD

Paul Neary, MD, FRCSI
Kenneth Palmer, MD
Eduardo Parra-Davila, MD, FACS
Ceana Nezhat, MD
Manuela Perez, MD, PhD
Cyril Perrenot, MD
Gary Poehling, MD
Vipul R. Patel, MD
Sonia L. Ramamoorthy, MD, FACS
Koon Ho Rha, MD, FACS, PhD
Judith Riess, PhD
Bernardo M. Rocco, MD
COL Robert Rush, MD
Richard Satava, MD, FACS
Brendan Sayers, MD
Daniel J. Scott, MD
Steve Schwaitzberg, MD
Neal Seymour, MD
Nazema Siddiqui, MD
Mika Sinanan, MD, PhD, FACS
Roger D. Smith, PhD
Hooman Soltanian, MD
Dimitrios Stefanidis, MD, PhD, FACS
Chandru Sundaram, MBBS
Robert Sweet, MD, FACS
Amir Szold, MD
Raju Thomas, MD
Oscar Traynor, MD
Edward Verrier, MD, FACS
Gregory S. Weinstein, MD
Thomas Whalen, MD

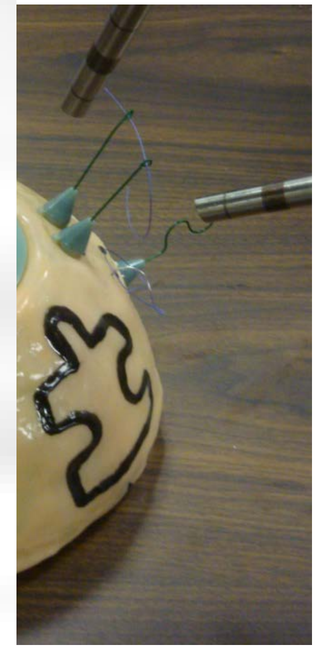
FRS Products

Or



F

ls Device



:com

Consensus Conference Process

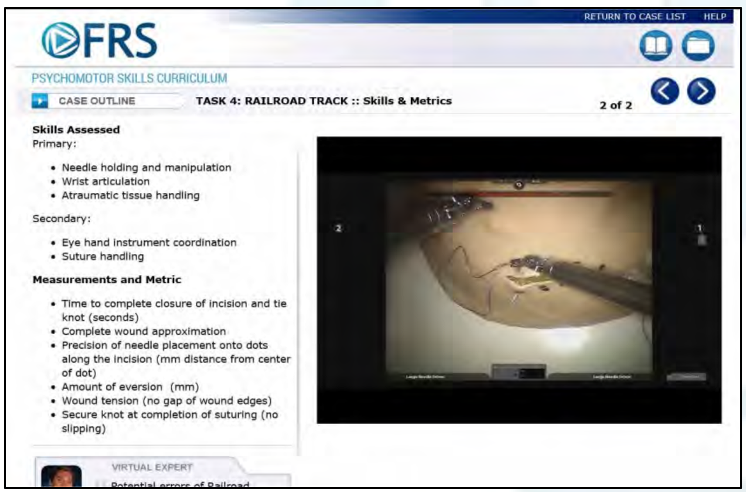
1. Outcomes Measures (Dec 12-13, 2011)
2. Curriculum Outline (April 29-30, 2012)
3. Curriculum Development (Aug 17-18, 2012)
4. Validation Criteria (November 17-18, 2012)
5. Validation Trials (Jan-Oct 2015)
6. Develop High Stakes Testing (2015)

FRS

Modules of the FRS Curriculum

Module 1: Introduction to Robotic Surgical Systems

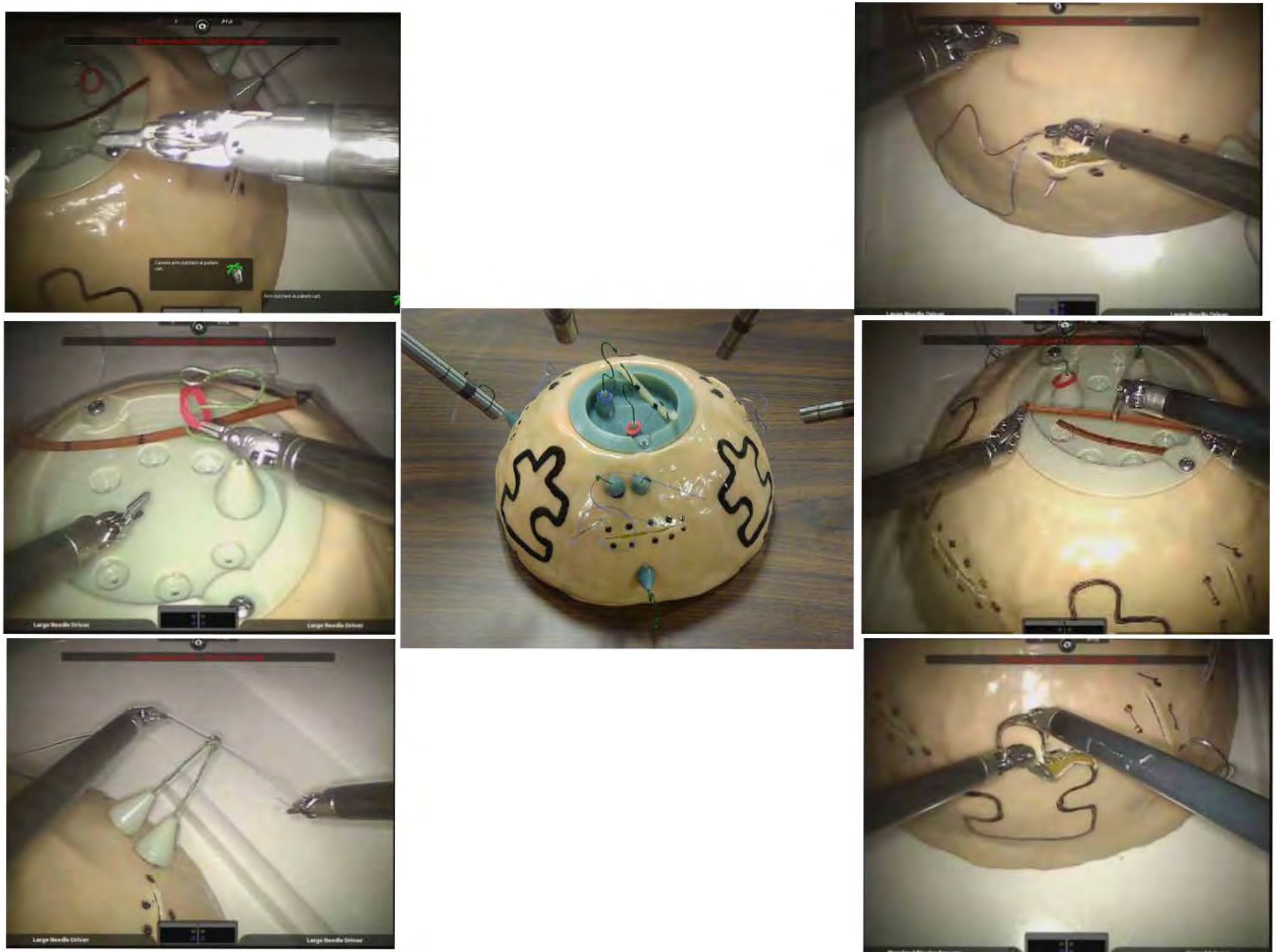
Module 2: Didactic Instructions



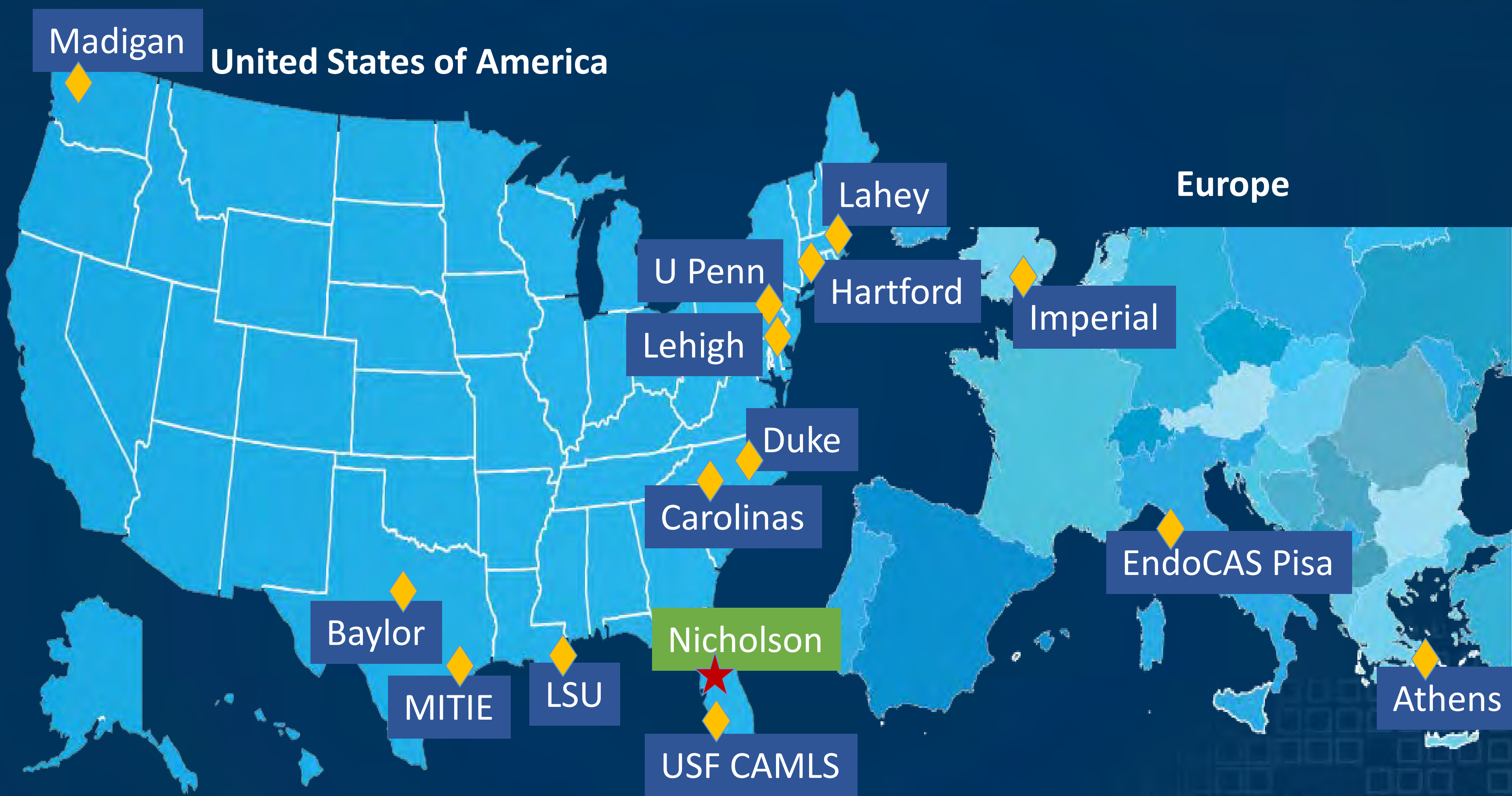
Module 3: Psychomotor Skills Curriculum

Module 4: Team Training and Communication Skills

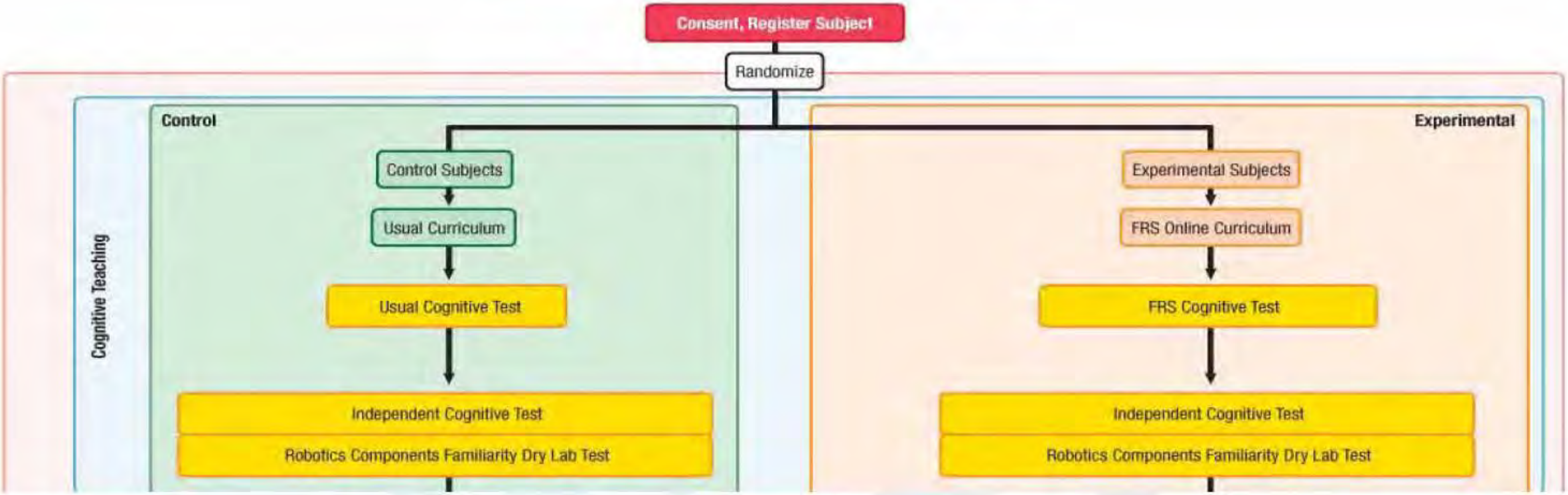
Prototype Process


Large Needle Driver
Large Needle Driver
Large Needle Driver
Large Needle Driver
Maryland Bipolar Forceps
Curved Scissors

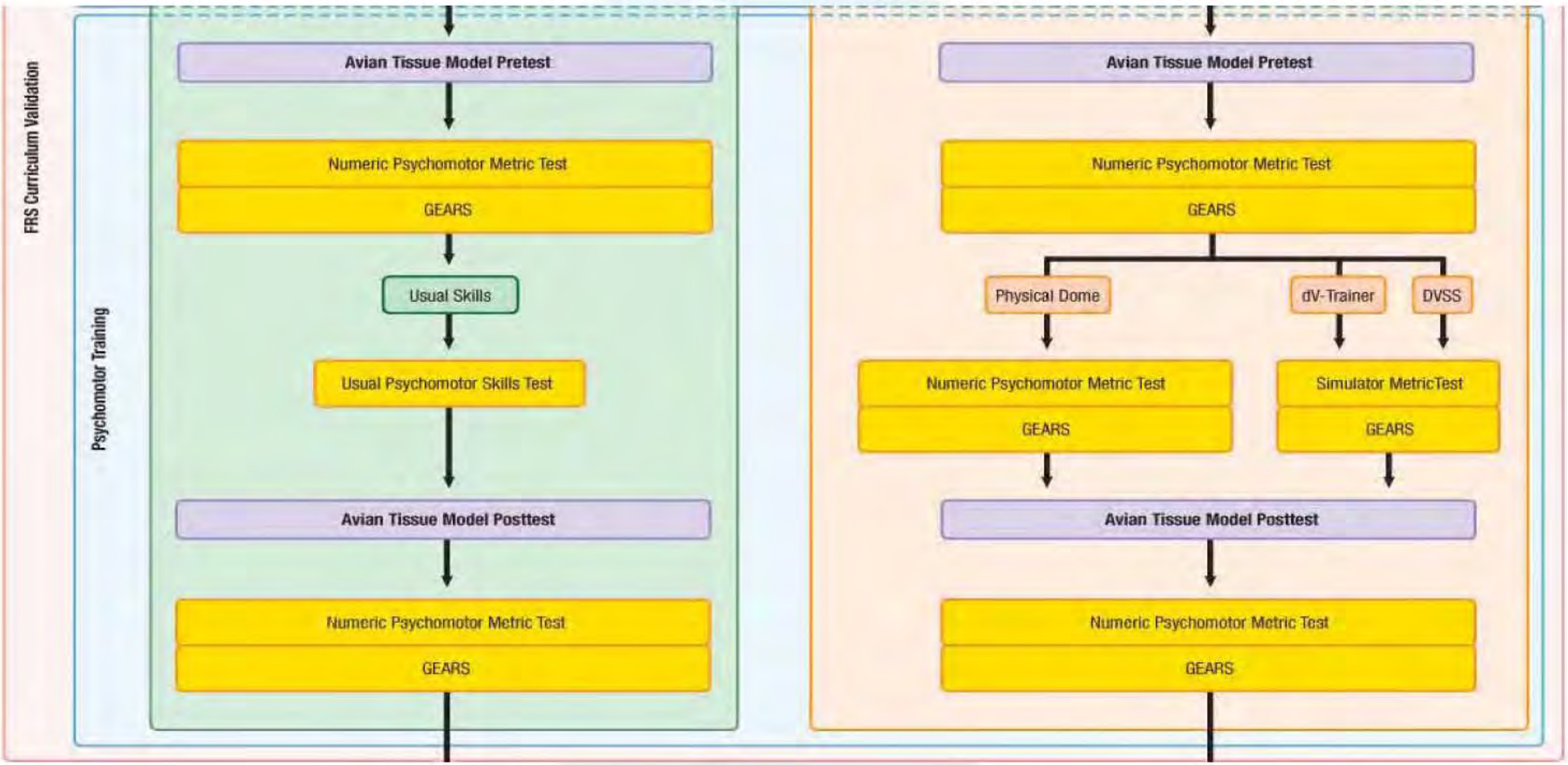
FRS Validation Sites



Validation Trial – Cognitive Learning

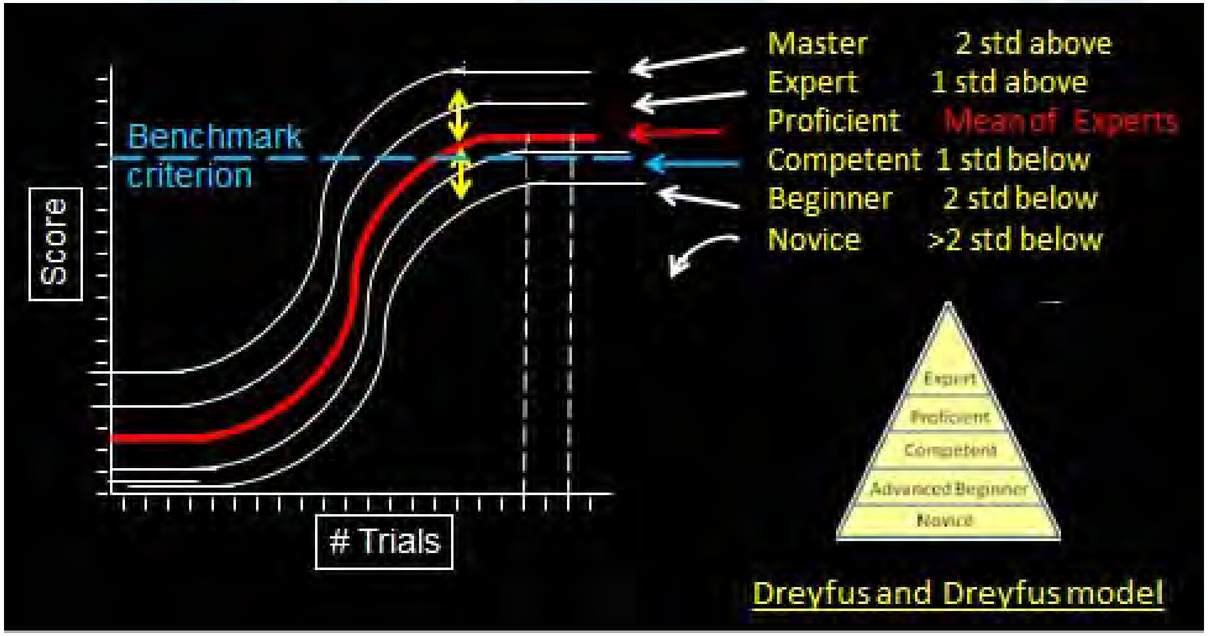


Validation Trial - Psychomotor



Setting Benchmarks

- Setting the Benchmark Criteria for Any Curriculum
- Train expert/experienced surgeons to their learning curve
- Two consecutive trials with no improvement
- Calculate the mean of their performance
- Calculate the standard deviation for other levels
- Proficiency is the Mean of Expert Performance



Summary

- Validation Trials are underway
  - Planned Completion in October 2015
- Online Curriculum and psychomotor device are openly available now
  - FRSurgery.org
  - FRSDome.com

Robotic Surgery Simulation

For every complex and expensive system there emerges a need for training devices and scenarios that will assist new learners in mastering the use of the device and understanding how to apply it with value. Intuitive Surgical's da Vinci robot is just such a system. It is currently the only FDA approved device for laparoscopic robotic surgery on human patients. Despite the 1.5 to 2 million dollar price tag, the device has seen rapid distribution and its implementation has led to the need to develop more efficient and effective training methods, as well as assessment and skill maintenance tools. In laparoscopic surgery, simulators have played an important role in improving the practice of surgery over the last 20 years [1,2]. The same trends and values will likely apply in robotic surgery.

The complexity, criticality, and cost associated with the application of the da Vinci surgical robot have stimulated the commercial creation of simulators which replicate the operations of this robot. There are currently four different simulation systems available for training and developing skills in robotic surgery: da Vinci Skills Simulator (Intuitive Surgical Inc.); dV-Trainer (Mimic Technologies, Inc.); RoSS (Simulated Surgical Skills LLC); and Robotix Mentor (Simbionix Ltd). Each of these possesses unique traits which make them valuable solutions for different types of users and learning environments. Investment in the simulators alone represents a major capital investment as costs range between $85,000 and $125,000 per device. Coupled with increases in direct costs from operating room time and supplies, robotic surgery implementation can have a profound impact on hospital finances [3,4,5,6,7]. These do not include costs associated with ensuring patient safety and surgeon training. From a business perspective, a hospital, college of medicine, robotic practice, or education center should calculate a financial return that will be achieved through the purchase and use of these devices.

Simulation and ROI

Rehearsal and simulation has been one of the primary means of developing and maintaining proficiency in specific skills for thousands of years. Lectures and written materials are the primary means for developing cognitive knowledge, but these are not effective at instilling psychomotor skills in any field. The skills needed in the hands, body, and coordination with the mind must be developed through practice. Rehearsal in a non-lethal environment has become the standard of practice for learning in military warfare and aviation. The lives of soldiers and pilots today are considered of significant value to demand a structured training program and measure of proficiency before entering a life threatening situation. Given the technology that has emerged in the late 20th and early 21st centuries, the means of rehearsal have shifted from drills to events that are mediated by computers with the ability to measure performance and provide constructive feedback to improve performance. Implementation of direct practice on simulators has led to significant quality improvements and a reduction in training costs to the tune of hundreds of millions of dollars. There is a 30 to 40 year history of success documented in the military literature, specifically regarding its return on investment. On average simulators cost only 5-20% of live training and can reduce the length of training by 10-30% [8].

Simulation devices are similarly a prominent tool in medical and surgical education. From the most primitive stuffed dolls with anatomical markings used in ancient Chinese societies, to the most current computer driven, three dimensional representations of living tissue, simulators have been helping to train surgeons for over a thousand years . While there are parallels between the military and medical fields regarding the risk to human life and the need for structured high fidelity training, there are also unique challenges in medicine. There are many confounders in medicine and the seemingly direct relationship between simulation and improved patient outcomes and reduction in costs is not robust. So, while most agree that simulation is a necessary requirement for surgical training, the lack of evidence coupled with the costs of devices and curricula development have led to a slower adoption of simulation in surgical education when compared to other industries. One of our goals is to use the lessons from our military and aviation industries in conjunction with published literature to identify key variables that impact the ROI of robotic surgical simulators and provide a loose framework for how simulators can be better utilized to achieve financial, educational, and patient safety benefits. We will focus on the concept of return on investment (ROI) for these simulators, not on the capabilities of any one specific device. Our goal is to assist the purchasers and users of these devices in determining whether such a device is a sound financial investment.

For a working copy of the ROI Model, capture this QR Code.

Or download from https://www.nicholsoncenter.com/our-research/

With this as a foundation, we used a systems analysis method popularized by Jay Forrester at Massachusetts Inst
ships between a large number of variables within the complex and dynamic hospital surgical program. This identi
changes in the training offered to robotic surgeons as well as longer-term variables which may take years to emer
those systems diagrams we distilled and grouped variables and relationships into the model shown in Figure 3. N
financial outcome. This model was useful in identifying the factors which contributed to ROI at a first, second, an
sions around the financial and non-financial variables that were changing due to the incorporation of simulation.

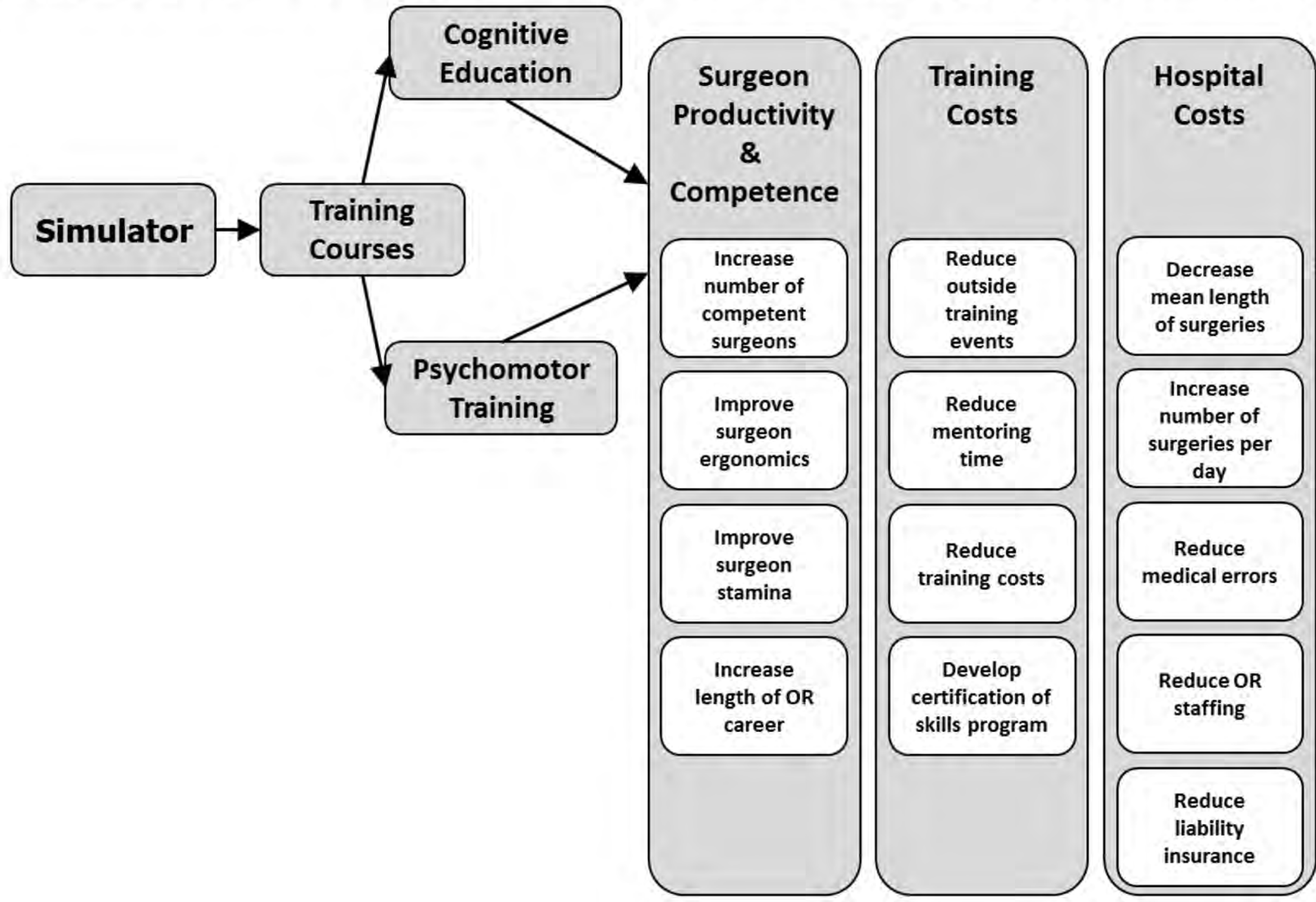


Figure 3. Effects of Simulation-based Training on Robotic Surgery E

Expert Opinion

Twelve expert interviews were conducted at several surgical and simulation centers around the country and inter
identifying and quantifying new variables. Surgical leaders interviewed included Randy Fagin MD, Texas Institute
Southwestern; Arnold Advincula MD, Celebration Health; Thomas Lendvay, MD, University of Washington College
nesota College of Medicine; Col Timothy Brand, MD, Madigan Army Medical Center; Jacques Hubert, MD, Univers
MD, Carolinas Healthcare System; Martin Martino, MD, Lehigh Valley Health Network; Mona Orady, MD, Clevelar
tems; and Michael Pitter, MD, Newark Beth Israel Medical Center.

There were several common themes identified from these expert interviews. Most agreed that having a defined
that it should be identified prior to purchasing a simulator. These robotic training curriculums consisted mostly o
combination of simulator and/or dry lab assessment. It was also clear that there were a variety of successful arra
simulation curriculum. The majority utilized a combination of the DVSS (Intuitive Surgical Inc.) and the dV-Traine
often a part of resident and fellow curriculums but approximately 50% of the sites also had attending level progra
cessibility of the equipment depending on its placement in the facility.

They verbalized a strong need for studies evaluating correlations between simulation and improved clinical outc
simulation training, and comparative studies between institutions with and without simulation centers. Curricul
well as remedial training following complications. The surgeons also showed a shared interest in the developme
plexity of the simulated exercises. Several believed that simulation could be used globally to decrease insurance

Video Game Experience and Basic Robotic Surgical Skills

Alyssa Tanaka M.S., Manuela Perez, M.D., Ph. D., Courtney Graddy M.H.A., & Roger Smith Ph.D.

FLORIDA HOSPITAL
NICHOLSON CENTER

BACKGROUND

- **PRIOR VIDEO GAME EXPERIENCE**

Rosser, Lynch, Cuddihy, Gentile, Klonsky, Merrell, 2007; Grantcharov, Bardram, Funch-Jensen, Rosenberg, 2003; Lehmann, Holmer, Gillen, Gröne, Zurbuchen, Ritz, & Buhr, 2013;

- **VIDEO GAME TRAINING**

Rosser, Gentile, Hanigan, & Danner, 2012; Badurdeen, Abdul-Samad, Story, Wilson, Down, & Harris 2010; Ju, Chang, Buckley, & Wang, 2012; Bokhari, Bollman-McGregor, Kahol, Smith, Feinstein, & Ferrara, 2010

- **FEW STUDIES HAVE LOOKED AT VIDEO GAMING IN ROBOTICS**

Chien, Suh, Park, Mukherjee, Oleynikov, & Siu, 2013; Harper, Kaiser, Ebrahimi,

PERCEPTUAL ABILITIES

BACKGROUND

VIDEO GAMES DEVELOP IMPORTANT PERCEPTUAL SKILLS

- **ACTION GAMES=VISUO-ATTENTION SKILLS**
- **STRATEGY GAMES=EXECUTIVE CONTROL SKILLS**
- **PUZZLE GAMES=PROBLEM-SOLVING/ RELATIONAL SKILLS**
- **CASUAL GAMES=HAND-EYE COORDINATION**

Green & Bavalier, 2003; Apperlay, 2006; Griffith, Voloschin, Gibb, & Bailey, 1983; Dorval & Pepin, 1986; Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1994

PURPOSE

COMPARE THE PERFORMANCE OF “EXPERT” VIDEO GAMERS TO “LAY PEOPLE,” MEDICAL STUDENTS, AND EXPERT ROBOTIC SURGEONS ON A ROBOTIC SURGERY SIMULATOR

FLORIDA HOSPITAL
NICHOLSON CENTER

METHODS

HYPOTHESES

H_1: VIDEO GAMERS WILL PERFORM BETTER THAN MEDICAL STUDENTS AND LAYPEOPLE

H_2: VIDEO GAMERS WILL DEMONSTRATE PERCEPTUAL SKILLS SIMILAR TO EXPERT SURGEONS

METHODS

VIDEO GAMERS:
DAILY USE ≥ 5 DAYS PER WEEK, ≥ 2 HOURS PER DAY

MEDICAL STUDENTS:
< 2 HOURS PER WEEK OF VIDEO GAME PRACTICE

LAY PEOPLE:
NO FORMAL MEDICAL OR VIDEO GAME EDUCATION

EXPERT ROBOTIC SURGEONS:
≥ 100 ROBOTIC CASES, ≥ 25 ROBOTIC CASES ANNUALLY

FLORIDA HOSPITAL
NICHOLSON CENTER

METHODS

DESIGN

Demographics Questionnaire

Cognitive Tests

Warm up
+
8 Trials of
2 Exercises

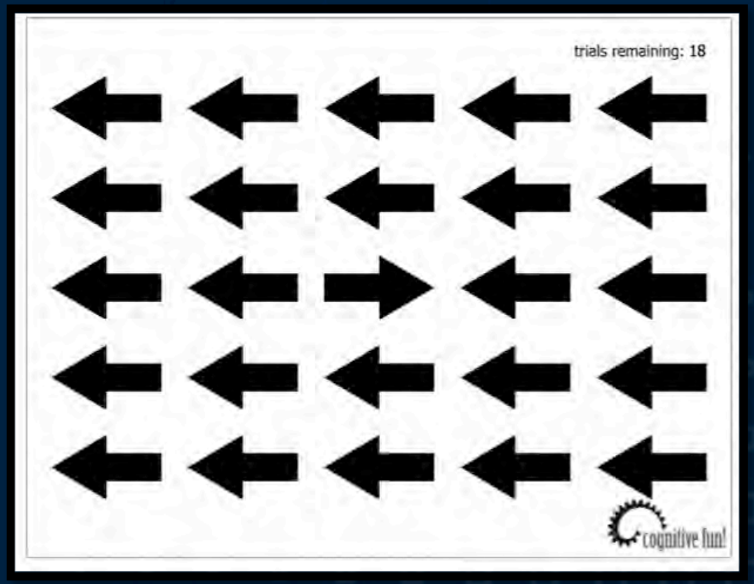
NASA TLX

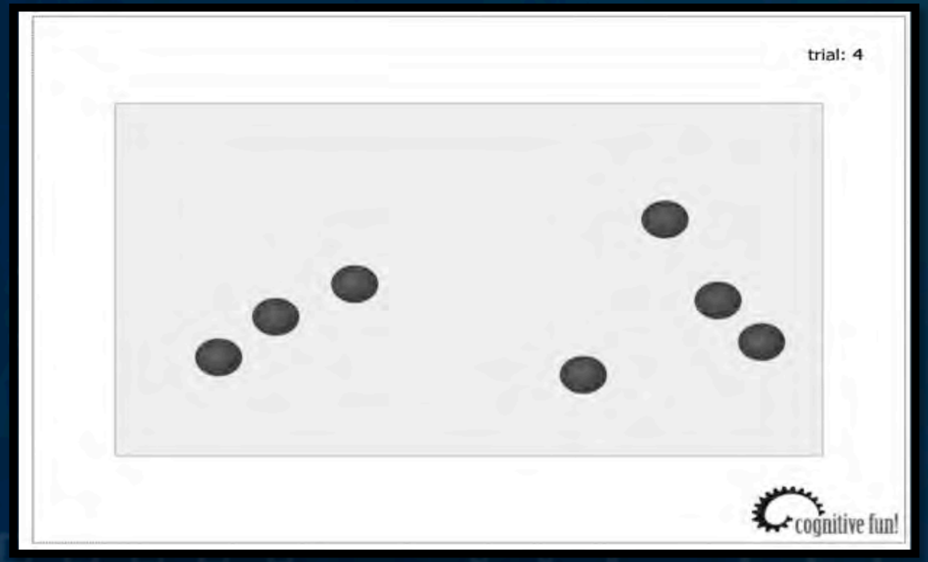
Post Questionnaire

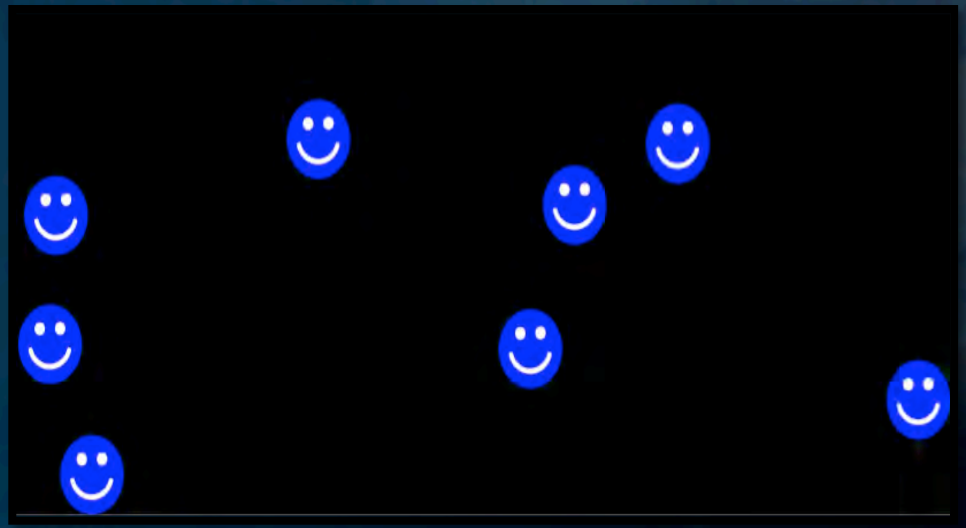
METHODS

COMPUTER-BASED PERCEPTUAL TESTS

- FLANKER
- FAST COUNTING
- MULTIPLE OBJECT TRACKING (MOT)





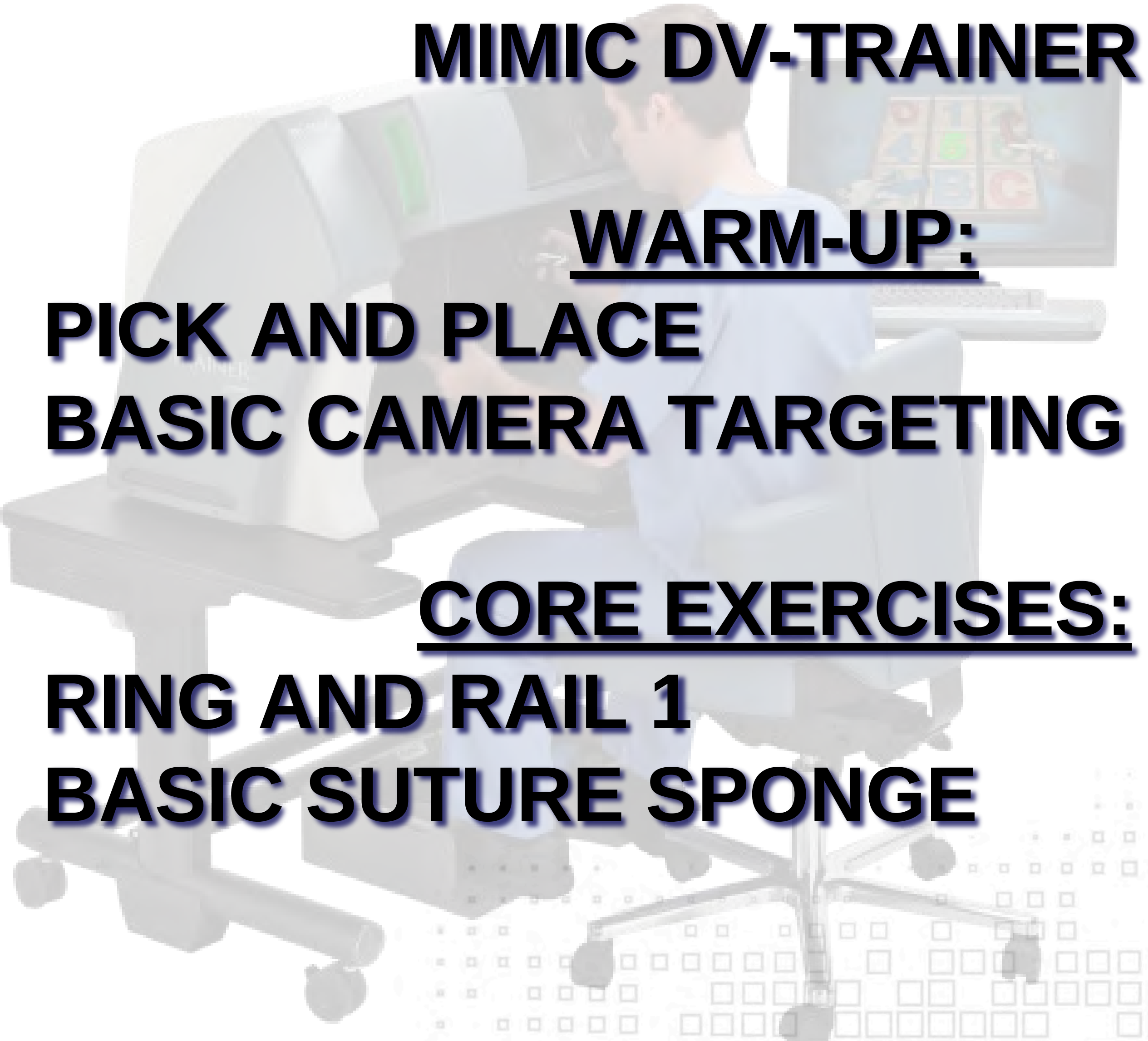


FLORIDA HOSPITAL
NICHOLSON CENTER

METHODS

SIMULATION

MIMIC DV-TRAINER



<u>WARM-UP:</u>

- **PICK AND PLACE**
- **BASIC CAMERA TARGETING**

<u>CORE EXERCISES:</u>

- **RING AND RAIL 1**
- **BASIC SUTURE SPONGE**

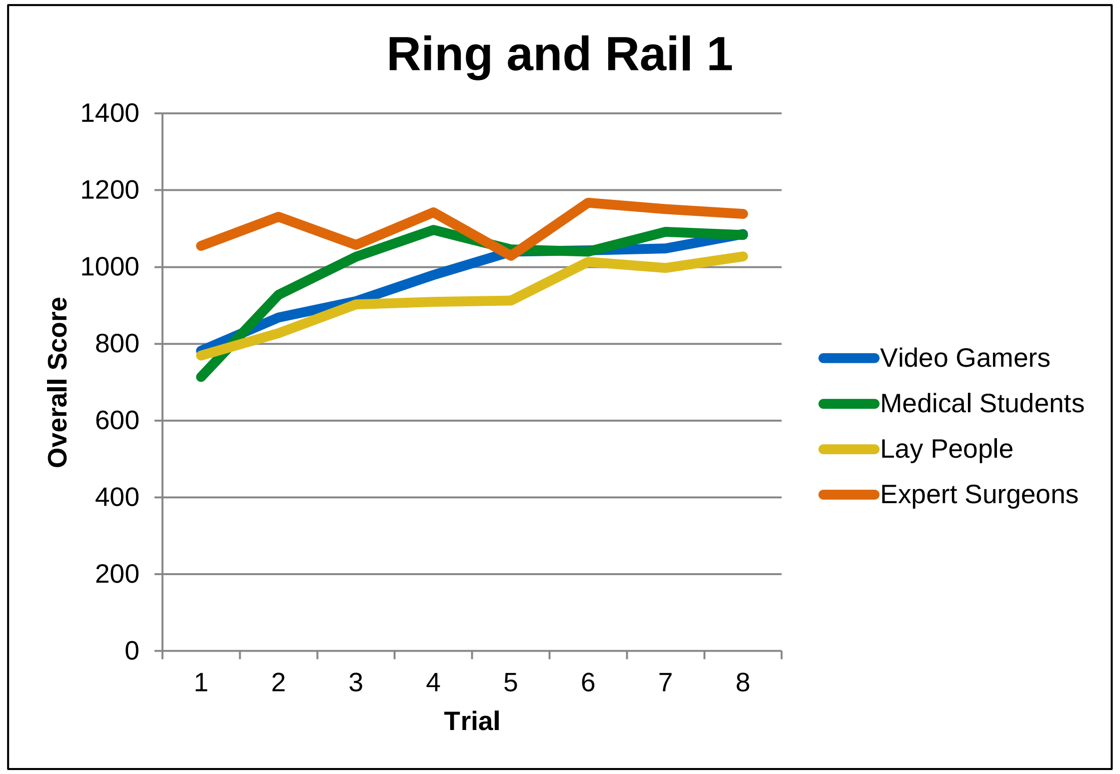
RESULTS

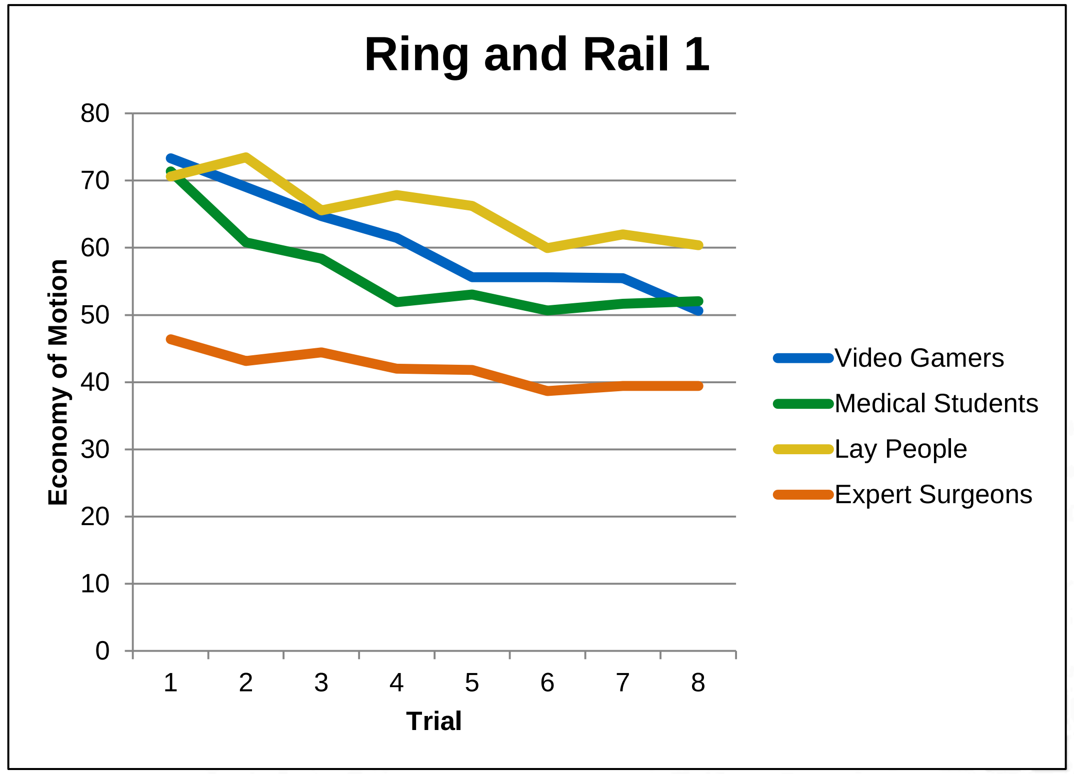
		Average Age	Male	Female	School Year	Average Hours Gaming per Week	Average Years of gaming	Total Robotic Cases	Cases per year
Gamers	n=40	25.38	77.50%	22.50%		11.71	17.85		
Medical Students	n=24	25.70	70.83%	29.17%	1st: 14 2nd: 3 3rd: 4 4th: 3				
Lay people	n=35	30.58	48.39%	51.61%					
Experts	n=6	43.33	83.33%	16.67%				503.33	127.5

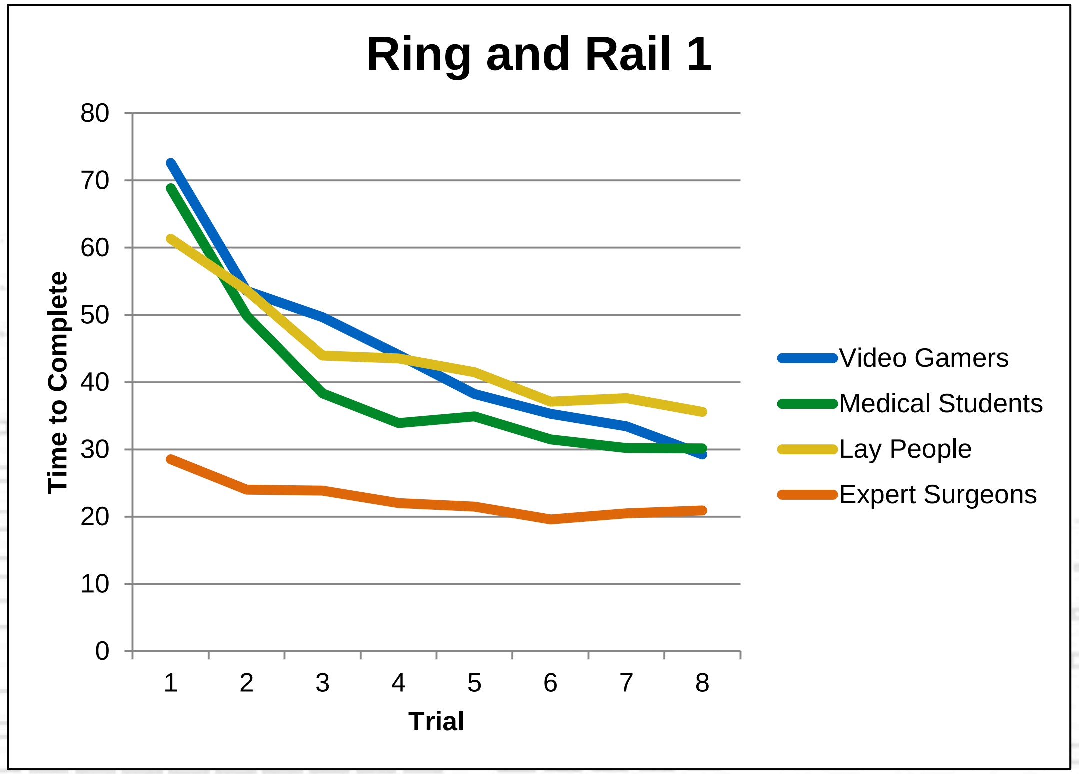
FLORIDA HOSPITAL
NICHOLSON CENTER

RESULTS

SIMULATION

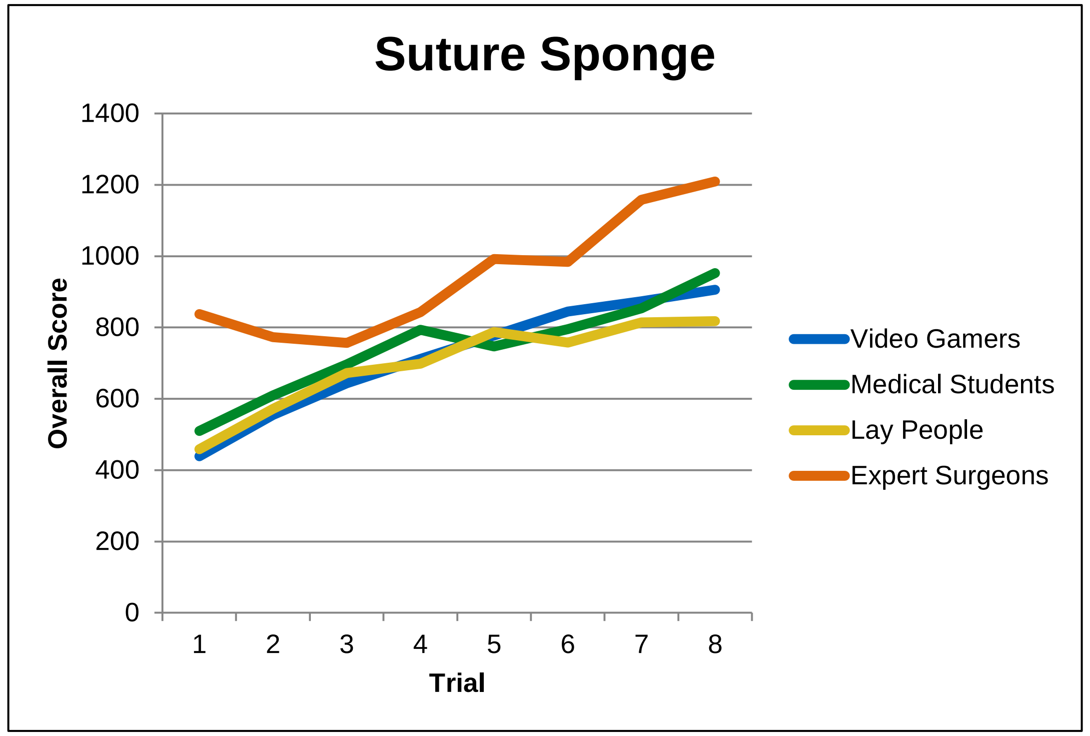


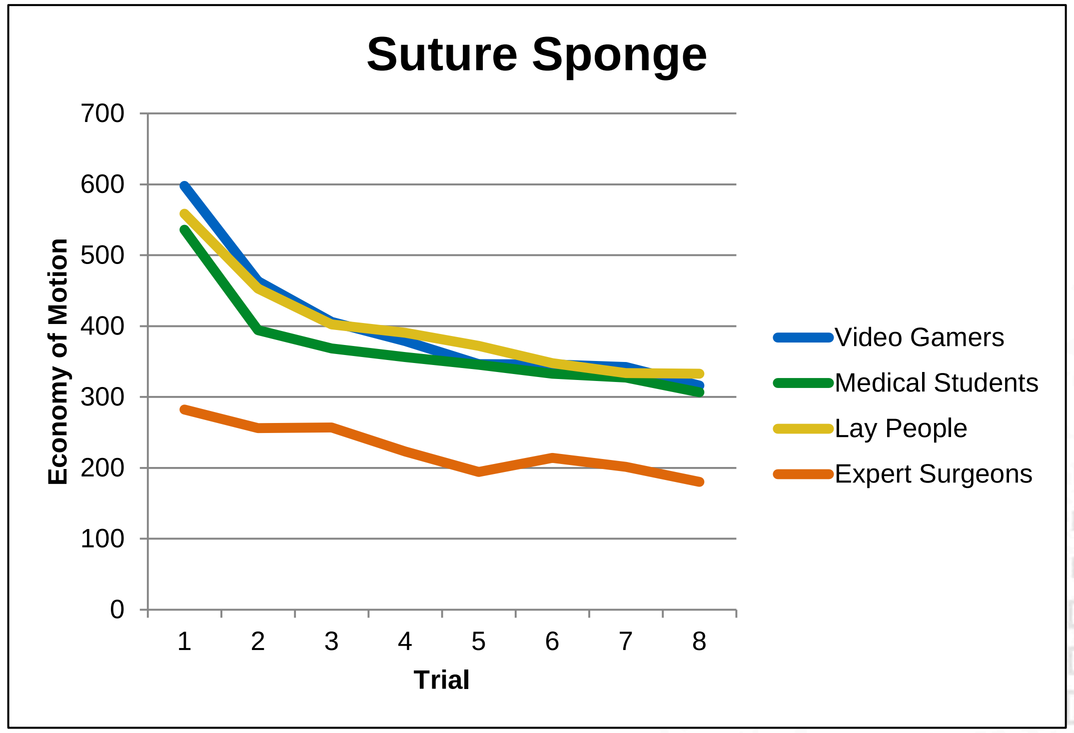


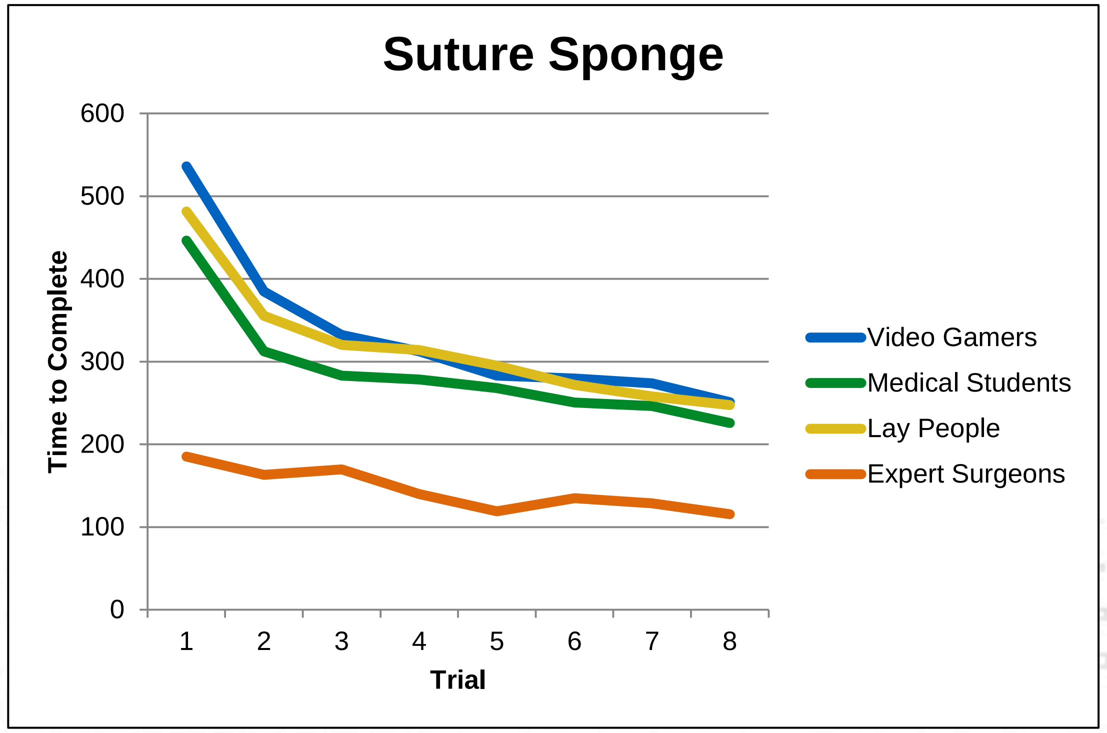


RESULTS

SIMULATION







Int J CARS
DOI 10.1007/s11548-015-1306-y

ORIGINAL ARTICLE

Impact of delay on telesurgical performance: study on the robotic simulator dV-Trainer

Manuela Perez[1,2,3,7] · Song Xu[1,4] · Sanket Chauhan[5,6] · Alyssa Tanaka[3] · Khara Simpson[3] · Haidar Abdul-Muhsin[3] · Roger Smith[3]

Received: 19 May 2015 / Accepted: 21 September 2015

Abstract

Purpose To determine the impact of communication latency on telesurgical performance using the robotic simulator dV-Trainer®

Methods Surgeons were enrolled during three robotic congresses. They were randomly assigned to a delay group (ranging from 100 to 1000 ms). Each group performed three times a set of four exercises on the simulator: the first attempt without delay (Base) and the last two attempts with delay (Warm-up and Test). The impact of different levels of latency was evaluated.

Results Thirty-seven surgeons were involved. The different latency groups achieved similar baseline performance with a mean task completion time of 207.2 s ($p > 0.05$). In the Test stage, the task duration increased gradually from 156.4 to 310.7 s as latency increased from 100 to 500 ms. In separate groups, the task duration deteriorated from Base for latency stages at delays ≥300 ms, and the errors increased at 500 ms and above ($p < 0.05$). The subjects' performance tended to improve from the Warm-up to the Test period. Few subjects completed the tasks with a delay higher than 700 ms.

Conclusion Gradually increasing latency has a growing impact on performances. Measurable deterioration of performance begins at 300 ms. Delays higher than 700 ms are difficult to manage especially in more complex tasks. Surgeons showed the potential to adapt to delay and may be trained to improve their telesurgical performance at lower-latency levels.

Keywords Telesurgery · Telemedicine/methods · Computer simulation · Robotic simulator · Internet

Grant Information: Funding Source: U.S. Army Telemedicine and Advanced Technology Research Center. Grant #: W81XWH-11-2-0158. The Grant was provided to the Florida Hospital Nicholson Center without personal benefit for the authors.

✉ Manuela Perez
m.perez@chu-nancy.fr

Roger Smith
roger.smith@flhosp.org

1 IADI Laboratory-INSERM-U947, Lorraine University, Allée du Morvan, 54500 Vandoeuvre-les-Nancy, France

2 General and Emergency Surgery Department, University Hospital of Nancy, Avenue du Maréchal de Lattre de Tassigny, 54035 Nancy, France

3 Florida Hospital Nicholson Center, 404 Celebration Place, Celebration, FL 34747, USA

4 Urology Department, University Hospital of Nancy, Allée du Morvan, 54511 Vandoeuvre-les-Nancy, France

5 Center for Evidence Based Simulation, Baylor University Medical Center, Dallas, TX, USA

6 Associate Professor Texas A&M Health Science Center, College Station, TX, USA

7 Hôpital Central Service de Chirurgie Générale et Urgences, Avenue du Maréchal de Lattre de Tassigny, 54035 Nancy, France

Abbreviations

ATM line	Asynchronous transfer mode line
ms	Millisecond
PB1	Peg-Board 1
CT2	Camera Targeting 2
TR1	Thread the Ring 1
ED1	Energy Dissection 1

Introduction

Robotic surgery was noted to be in its infancy in 2004, [1] but now this advanced technology is on its way to young adulthood [2]. It has become a standard in complex surgery [3]. The mature experience will likely include the achievement of remote telesurgery, a future challenge for robotic surgeons [4,5].

The first transatlantic human telesurgery procedure was performed in 2001 [6]. Since the proof of concept, telesurgery remains a complex and uncommon process that holds promise in overcoming challenging situations (remote medicine for underserved regions, surgery in the battlefield, surgery in space, etc.) [7,8]. Many teams have worked on the telesurgery process and tried to achieve remote telesurgery procedures using available technical resources for the video flux transfer [7,9,10]. In telesurgery, the control signal sent from the master console is transferred over a network to the robot arms followed by a corresponding movement of the surgical instruments. The video images are then returned to the surgeon site. The data transmission requires an encoding, transmission, and decoding process in which a time delay, or latency, is inevitably produced. Latency is correlated with the amount of data and the quality of network. The first transatlantic human telesurgery (with the Zeus robot) used sophisticated dedicated asynchronous transfer mode (ATM) lines with a transmission delay around 150 ms [6]. Dedicated lines, however, are not always feasible in routine clinical situations. The public Internet bridging the world could be an easy and accessible resource to transmit this data. Even so, the network availability would be at the price of increasing latency measured approximately 450–900 ms [11].

It would be valuable to clarify the impact of the latency on surgical performances before future implementations of telesurgery. Two thresholds need to be established: The first is the smallest latency that can be detected by surgeons which will influence their performance, and the second is the level of latency that makes the surgery unsafe. Unsafe surgery is associated with an increase in errors. A previous study on this topic highlighted the impact of delay on performance degradation using the dV-Trainer®. The authors evaluated the effects of delay varying between 100 and 1000 ms, and found that latencies ≤300 ms had a small impact on performance. Subjective evaluation then suggested that surgery became quite difficult at delays ≥800 ms [12]. However, this study only included medical students as the subjects. Additional experiments should be performed with experienced surgeons, especially those experienced with robotic systems which would be needed to implement telesurgical procedures.

The present study aims to evaluate, on a surgeon population, the impact of different latency levels on performances in four simulated robotic tasks.

Material and methods

Exercises and subjects

We designed a prospective, observational study conducted on the robotic surgical simulator dV-Trainer® (Mimic technologies Inc., Seattle, USA). This tool has demonstrated face, content, construct, and concurrent validity in previous studies [13,14]. Based on expert opinion and literature review [14,15], we chose four exercises for the test that would be performed in a constant easy-to-difficult order: (a) Peg-Board 1 (PB1)—pick up and transfer rings sequentially from the Peg-Board to a single peg on the floor; (b) Camera Targeting 2 (CT2)—manipulate the camera to precisely focus and zoom on a target sphere; pick up and move a stone into a designated basket; (c) Thread the Rings 1 (TR1)—pass a needle and suture through a number of flexible eyelets; (d) Energy Dissection 1 (ED1)—isolate a large blood vessel by cauterizing and cutting small branching blood vessels that anchor the large vessel (Fig. 1). Both basic (endowrist manipulation, camera control, clutching) and challenging (suturing, dissection) skills were covered with these exercises. The dV-Trainer® simulator permitted us to introduce fixed latencies into the exercises between the gesture on the grips and the visual feedback on the console.

After institutional review board approval, we recruited subjects—fellows and attending surgeons—during three robotic surgery conferences. All the experiments involving human participants were in accordance with the ethical standards of the institutional research committee, as well as the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained [3] from all individual participants.

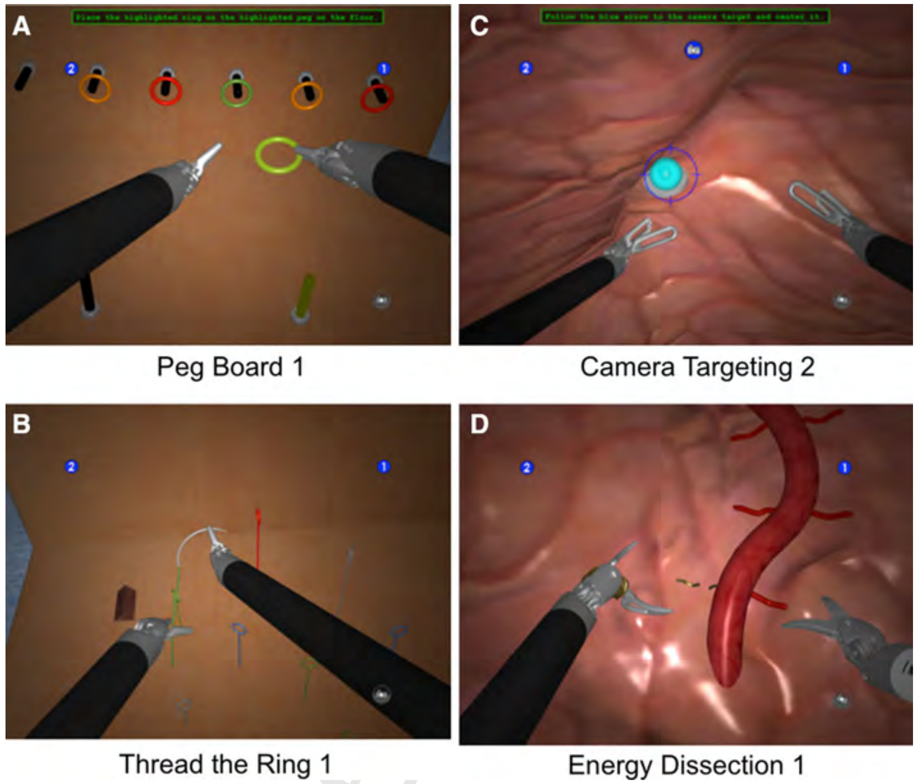
Procedures

Each participant received a unique identification number under which all his/her data would be collected, and then completed a questionnaire concerning demographic data (including surgical experience and related activities).

Each subject was randomly and blindly assigned a latency varying between 100 and 1000 ms with increments of 100 ms. Before the trials on dV-Trainer®, they received standard instruction on its use in a familiarization period. After that, they performed all four exercises in order without delay (Base). The results provided their baseline performance. Then they repeated the same set of exercises twice with the assigned latency (Warm-up and Test). The Warm-up period allowed them to become familiar [4] with latency and to acquire short-term adaptation (Fig. 2).

Journal: 11548 MS: 1306 ☐ TYPESET ☐ DISK ☐ LE ☐ CP Disp.:2015/9/25 Pages: 7 Layout: Large

Fig. 1 The four dV-Trainer® exercises: Peg-Board 1 (**a**), Camera Targeting 2 (**b**), Thread the Ring 1 (**c**), and Energy Dissection 1 (**d**)



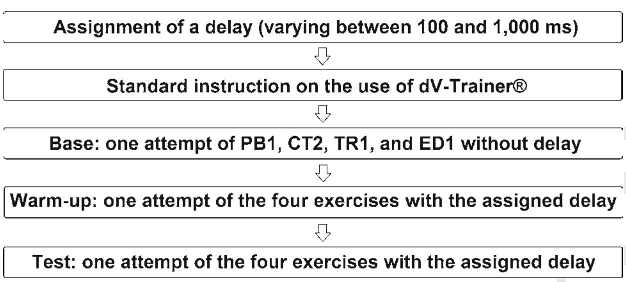


Fig. 2 Experimental procedures

Metrics

The dV-trainer includes a built-in scoring system. The values of the following metrics were automatically recorded after each exercise: time to complete the exercise (in seconds), instrument motion (in centimeters), master workspace range (in centimeters), excessive instrument force (in seconds), instruments out of view (in centimeters), instrument collisions, drops, etc. An overall score representing a combination of these criteria was also automatically generated.

Based on our experience, the task completion time is the most sensitive and reliable measure to the impact of delay [12]. We thus chose this measure to represent the results. In addition, the mean score of all error metrics was calculated in order to evaluate the latency impact on errors.

Statistics

Data were analyzed using the R statistical software. A repeated-measures ANOVA (mixed-effects model) was used to determine the differences in performances between various latency groups (with FDR p value correction), and also between the three periods in each latency group (with Holm correction). Statistical significance was determined at $p < 0.05$.

Results

Complete data

Final data were derived from 37 surgeons. Twenty-three persons had robotic experience, with an average of 2.7 years (ranging from 1 to 9 years). All subjects completed the three stages from Base to Test, but some of them did not complete all the exercises. For example, four subjects were included in the 100 ms group, but one of them did not complete the exercises of CT2 and TR1. The groups from 700 to 1000 ms were combined due to the limited subject number (Table 1).

Results across exercises

The different latency groups achieved similar baseline performance with a mean task completion time of 207.2 s ($p >$

Journal: **11548** MS: **1306** TYPESET DISK LE CP Disp.:**2015/9/25** Pages: **7** Layout: **Large**

Table 1 Demographic data

	100 ms	200 ms	300 ms	400 ms	500 ms	600 ms	700–1000 ms
n	4	8	2	7	4	7	5
Complete PB1	4	7	2	5	4	6	3
Complete CT2	3	6	2	5	4	4	5
Complete TR1	3	7	2	6	4	1[a]	1[a]
Complete ED1	4	7	2	7	3	3	1[a]
Age (years)	38.5 ± 7.0	45.4±10.5	48.5±13.4	47.4±12.5	37.8±4.3	44.6±8.4	42.6±9.4
Position (*n*)	Fellow (2) attending (2)	Fellow (3) attending (5)	Fellow (0) attending (2)	Fellow (1) attending (6)	Fellow (1) attending (3)	Fellow (0) attending (7)	Fellow (1) attending (4)
Laparoscopic experience (years)	6.3±1.9	11.8±7.9	10.5±3.5	16.7±13.1	7.3±1.5	13.6±6.6	14.0±8.4
Robotic experience (years, median)	0	0	1.25	1	0.5	2	3

[a] Data were not used due to the single number

0.05). An increasing tendency of the task duration with delay was observed in the two latency stages. In the Test period, the mean task duration increased from 156.4 s at 100 ms to 310.7 s at 500 ms. When comparing this measure between any two latency groups, statistical significance was achieved in the comparisons of the 100 ms group versus the 400 and 500 ms groups ($p < 0.05$; Fig. 3).

Subjects demonstrated the tendency to improve their performances from the Warm-up to the Test period. The task completion time deteriorated from the baseline to the two latency stages at 300 ms and above, although statistical significance was not achieved at 300 ms due to the limited subject number (Fig. 3). The comparison results between the three periods in each latency group are illustrated in Fig. 4.

The mean error score deteriorated from baseline to latency stages at 500 ms and above ($p < 0.05$). For example, in

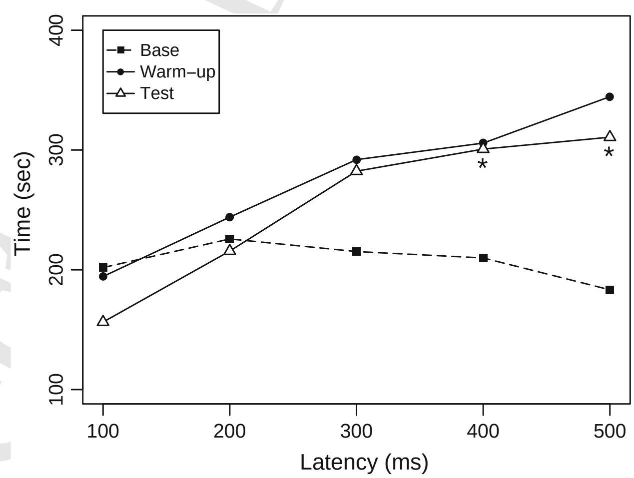


Fig. 3 The mean task completion time across the four test exercises in each latency group. * Difference was determined compared to the 100 ms group ($p < 0.05$). The groups of 600 and 700–1000 ms were not included due to insufficient data in certain exercises

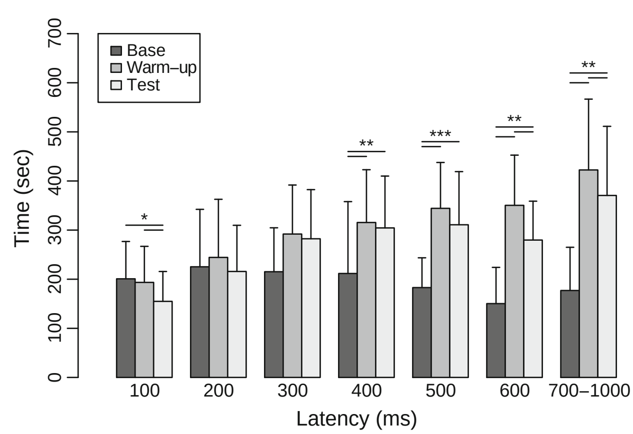


Fig. 4 Comparisons of the task completion time between the three periods in each latency group (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). The group 600 ms includes only the results across PB1, CT2, and ED1; the group 700–1000 ms includes the results across PB1 and CT2

Journal: 11548 MS: 1306 ☐ TYPESET ☐ DISK ☐ LE ☐ CP Disp.:2015/9/25 Pages: 7 Layout: Large

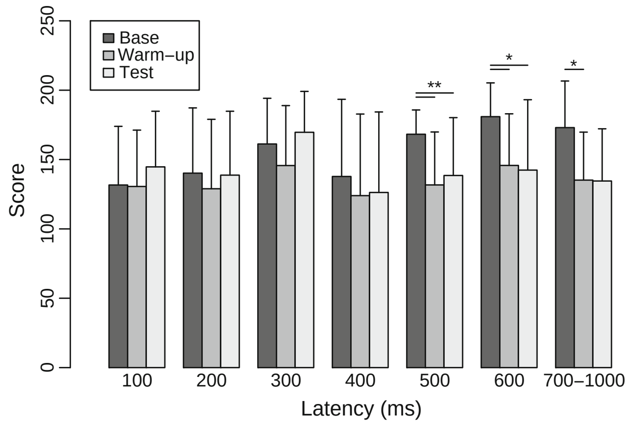


Fig. 5 Comparisons of the mean error score between the three periods in each latency group (* $p < 0.05$; ** $p < 0.01$). The group 600 ms includes only the results across PB1, CT2, and ED1; the group 700–1000 ms includes the results across PB1 and CT2

500 ms group, the score decreased from 168.2 (out of 200) to 138.5 from the Base to the Test period (Fig. 5).

Results in separate exercises

An increasing tendency of the task completion time with latency was observed in the two latency periods of the four exercises. The degradation of performances between baseline and latency stages started with 300, 500, 100, and 300 ms in PB1, CT2, TR1, and ED1, respectively (Fig. 6).

Incomplete data

Eighty incomplete exercises in latency stages derived from 26 subjects were identified. They included 18 PB1, 18 CT2, 26 TR1, and 18 ED1. Subjects were physically unable to complete these delayed exercises. Fifty-three (66.25 %) exercises were stopped by the subjects at a mean time of 9.8 min (586.01 ± 14.54 s). The ratio of incomplete exercises was relatively higher in high-delay groups (Fig. 7).

Discussion

We aimed to determine the latency effects on surgical performances in experienced surgeons who are unfamiliar with latency and the simulator device, to establish the threshold delays in telesurgery. Overall, the gradually increasing latency has an increasing impact on performances, and the performance deterioration consistently begins at 300ms. Latencies of 100 and 200 ms seemed to have no clear effect, and the 100 ms group had improving performance from the Base to the Test stage. This improvement likely corresponds to the learning effects of basic simulator manipulation and

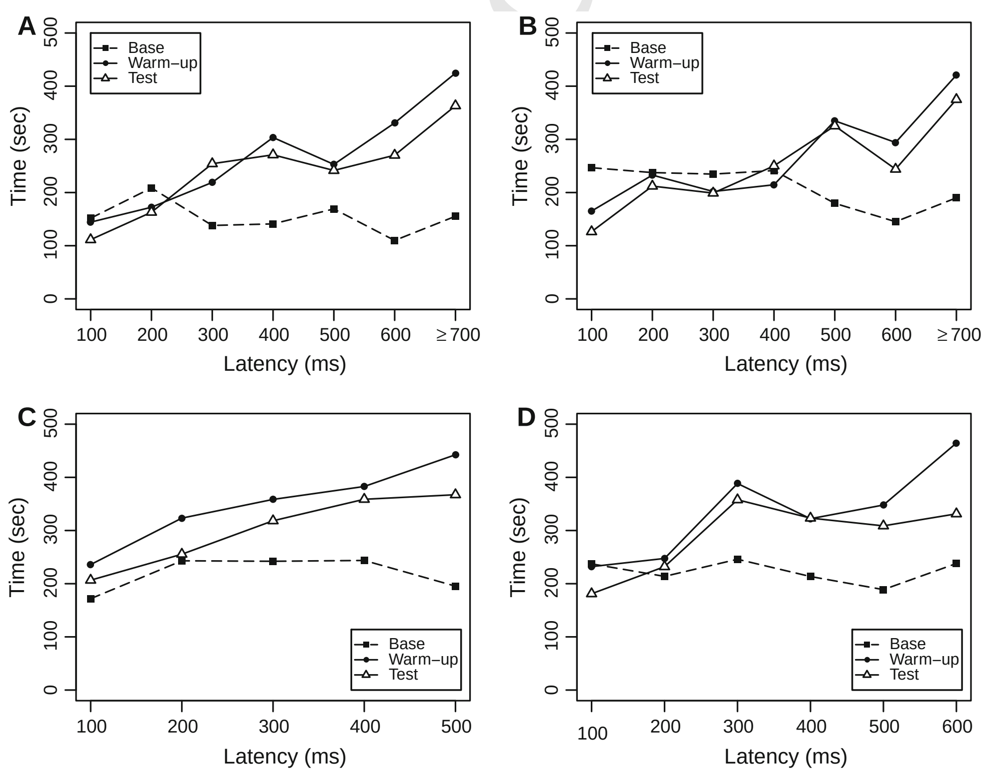


Fig. 6 The mean task completion time in each latency group of the four exercises: Peg-Board 1 (**a**), Camera Targeting 2 (**b**), Thread the Ring 1 (**c**), and Energy Dissection 1 (**d**)

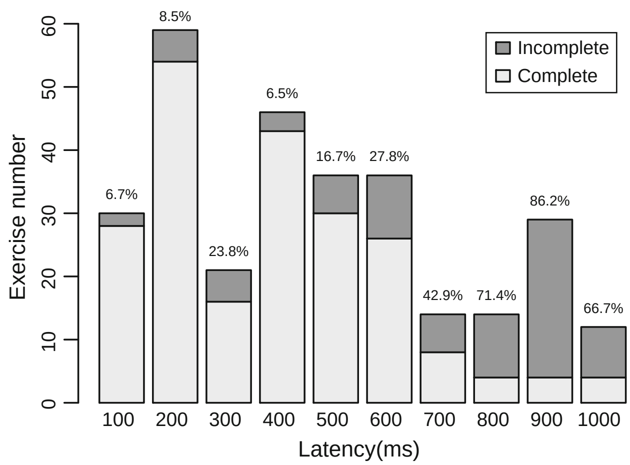


Fig. 7 The numbers of complete and identified incomplete exercises at each latency level. The numbers above the *bars* represent the percentage of incomplete exercises

further proves that 100 ms does not have a significant influence. For the superior threshold, delays equal or higher than 700 ms seem to be difficult to manage especially in complex tasks. Only one subject was able to complete the tasks at 700 ms, and only the easiest exercises (PB1, CT2) were finished at 800–1000 m. In the previous study with trained medical students, the similar threshold was highlighted and the authors suggested telementoring as a safer choice [12]. Telementoring is an application of telemedicine that involves the remote guidance of a procedure when the local operator has limited experience with the technique [16]. However, in this study, the error rate significantly increased from non-latency to latency stages at delays ≥500 ms, which may indicate an increase in surgical risk. We would consider this value as the superior threshold, and telesurgery should not be recommended in this condition for most surgeons [17]. This does not mean that procedures cannot be performed at higher-latency levels, and results could be better for experienced robotic surgeons, especially when given an opportunity to rehearse in an environment including latency, such as with a simulator. Current research is still limited, and outcome data are lacking to demonstrate the feasibility and safety of telesurgery with high delays. In a previous published study, a nephrectomy was performed on a swine under a delay of 900 ms. Two surgeons performed the procedure, one in the remote site console and the other in the local site console [11]. In this article, no outcome data were provided, such as surgical performance and the mental stress of surgeons.

Surgeons have been shown to have the potentials to adapt to delays [18]. Similar tendency was also observed in our study: Performances improved from Warm-up to Test. It suggests that surgeons may be trained on latency to improve their telesurgical performance. However, the improvement observed here is not clearly attributable to adaptation through experience with latency. It may also be the result of improvements in psychomotor simulator manipulation. Despite the overall tendency across exercises, our results also demonstrate that the impact of latency is related to the difficulty of procedures. Latency affected performances on different levels for the four chosen exercises: The performance deterioration started at a high delay (500 ms) for the simple exercise CT2 and at a low level (100 ms) in the more challenging TR1. This fact indicates that the minimum influential and the maximum acceptable delays could be different in surgical procedures with different complexity.

For the challenging exercises that may better represent real surgical scenarios, we have chosen TR1 and ED1 instead of the more complex exercises like “Suture Sponge” or “Tubes.” This is because many surgeons were not sufficiently familiar (or proficient) with the robot or the simulator. In this study, few tasks were completed at delays higher than 700 ms. One might anticipate that the results would be even worse if applying more challenging exercises.

Participants have demonstrated the efforts to complete the tasks even with considerable latencies. In the identified 80 incomplete exercises, only a few subjects terminated their participation soon after beginning. The mean duration of attempt was 7.5 min per exercise. This effort could minimize the bias of experiments. It is also interesting to observe that many persons stopped at about 10 min. It seems that this is a threshold beyond which surgeons could no longer endure the effects of latency.

This study has potential limitations: Although we recruited more than 60 surgeons, the final completion rate was lower than expected. The small number of subjects in each latency group is a shortcoming of the study. We did not merge different latency groups because the objective was to evaluate the impact of each latency level, and an interval of 100 ms may already cause difference. Also the distribution of subjects was not equivalent in different latency groups, primarily due to subjects choosing to terminate their participation before completing the entire experiment. Fewer subjects were included in the 300 ms group. Moreover, many surgeons failed to complete the tasks at high delays due to the difficulty of manipulation under these conditions. In addition, all subjects were novices in telesurgery (or latencies) since this technology is currently only available in research settings.

A complementary study will be necessary to assess the performance degradation induced by latency on robotic surgery experts, and to investigate whether latency training could be used to overcome the challenges of telesurgery.

Conclusion

This study was conducted on surgeons with limited experience using the dV-Trainer simulator, and the results demon-

Journal: 11548 MS: 1306 TYPESET DISK LE CP Disp.:2015/9/25 Pages: 7 Layout: Large

strated that performances (time to perform, score, error) deteriorate gradually as latency increases. The impact of delay is related to the difficulty of the procedures, but overall, delays of 100 to 200 ms have no significant impact, and a delay higher than 500 ms causes a noticeable increase in surgical risk. Surgery becomes extremely difficult and should be avoided at delays higher than 700 ms. Telementoring could be an option in this situation. Surgeons have the potential to adapt to latency, and they may be trained to improve their telesurgical performances using devices like simulators of robotic systems.

Compliance with ethical standard

Conflict of interest The authors Manuela Perez, Song Xu, Sanket Chauhan, Alyssa Tanaka, Khara Simpson, Haidar Abdul-Muhsin and Roger Smith declare they have no disclosure or conflict of interest to declare

References

1. Lanfranco AR, Castellanos AE, Desai JP, Meyers WC (2004) Robotic surgery: a current perspective. Ann Surg 239:14–21. doi:10.1097/01.sla.0000103020.19595.7d
2. Kim SP, Boorjian SA, Shah ND, Weight CJ, Tilburt JC, Han LC, Thompson RH, Trinh QD, Sun M, Moriarty JP, Karnes RJ (2013) Disparities in access to hospitals with robotic surgery for patients with prostate cancer undergoing radical prostatectomy. J Urol 189:514–520. doi:10.1016/j.juro.2012.09.033
3. Ghani KR, Sukumar S, Sammon JD, Rogers CG, Trinh QD, Menon M (2014) Practice patterns and outcomes of open and minimally invasive partial nephrectomy since the introduction of robotic partial nephrectomy: results from the nationwide inpatient sample. J Urol 191:907–912. doi:10.1016/j.juro.2013.10.099
4. Haidegger T, Sándor J, Benyó Z (2011) Surgery in space: the future of robotic telesurgery. Surg Endosc 25:681–690. doi:10.1007/s00464-010-1243-3
5. Gambadauro P, Torrejón R (2013) The "tele" factor in surgery today and tomorrow: implications for surgical training and education. Surg Today 43:115–122. doi:10.1007/s00595-012-0267-9
6. Marescaux J, Leroy J, Gagner M, Rubino F, Mutter D, Vix M, Butner SE, Smith MK (2001) Transatlantic robot-assisted telesurgery. Nature 413:379–380. doi:10.1038/35096636
7. Anvari M, McKinley C, Stein H (2005) Establishment of the world's first telerobotic remote surgical service: for provision of advanced laparoscopic surgery in a rural community. Ann Surg 241:460–464
8. Anvari M (2007) Telesurgery: remote knowledge translation in clinical surgery. World J Surg 31:1545–1550. doi:10.1007/s00268-007-9076-5
9. Rayman R, Croome K, Galbraith N, CcCulure R, Morady R, Peterson S, Smith S, Subotic V, Van Wynsberghe A, Primak S (2006) Long-distance robotic telesurgery: a feasibility study for care in remote environments. Int J Med Robot Comput Assist Surg MRCAS 2:216–224. doi:10.1002/rcs.99
10. Rayman R, Croome K, Galbraith N, CcCulure R, Morady R, Peterson S, Smith S, Subotic V, Van Wynsberghe A, Pate R, Primak S (2007) Robotic telesurgery: a real-world comparison of ground- and satellite-based internet performance. Int J Med Robot 3:111–116. doi:10.1002/rcs.133
11. Sterbis JR, Hanly EJ, Herman BC, Marohn MR, Broderick TJ, Shih SP, Harnett B, Doarn C, Schenkman NS (2008) Transcontinental telesurgical nephrectomy using the da Vinci robot in a porcine model. Urology 71:971–973. doi:10.1016/j.urology.2007.11.027
12. Xu S, Perez M, Yang K, Perrenot C, Felblinger J, Hubert J (2014) Determination of the latency effects on surgical performance and the acceptable latency levels in telesurgery using the dV-Trainer(®) simulator. Surg Endosc. doi:10.1007/s00464-014-3504-z
13. Lee JY, Mucksavage P, Kerbl DC, Huynh VB, EtafyM Mc, Dougall EM (2012) Validation study of a virtual reality robotic simulator-role as an assessment tool? J Urol 187:998–1002. doi:10.1016/j.juro.2011.10.160
14. Perrenot C, Perez M, Tran N, Jehl JP, Felblinger J, Bresler L, Hubert J (2012) The virtual reality simulator dV-Trainer(®) is a valid assessment tool for robotic surgical skills. Surg Endosc 26:2587–2593. doi:10.1007/s00464-012-2237-0
15. Ruparel RK, Taylor AS, Patel J, Patel V, Heckman MG, Rawal B, Leveillee RJ, Thiel DD (2014) Assessment of virtual reality robotic simulation performance by urology resident trainees. J Surg Educ 71:302–308. doi:10.1016/j.jsurg.2013.09.009
16. Rosser JC Jr, Gabriel N, Herman B, Murayama M (2001) Telementoring and teleproctoring. World J Surg 25:1438–1448
17. Lendvay TS, Hannaford B, Satava RM (2013) Future of robotic surgery. Cancer J Sudbury Mass 19:109–119. doi:10.1097/PPO.0b013e31828bf822
18. Nguan CY, Morady R, Wang C, Harrison D, Browning D, Rayman R, Luke PPW (2008) Robotic pyeloplasty using internet protocol and satellite network-based telesurgery. Int J Med Robot 4:10–14. doi:10.1002/rcs.173

Journal: **11548** MS: **1306** ☐ TYPESET ☐ DISK ☐ LE ☐ CP Disp.:**2015/9/25** Pages: **7** Layout: **Large**

Journal: 11548
Article: 1306

Springer
the language of science

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details required	Author's response
1.	Please check and confirm that the authors and their respective affiliations have been correctly identified and amend if necessary.	No error on the author' affiliations
2.	Please check if the corresponding author is correctly identified.	Corresponding author correctly identified
3.	Please confirm the section headings are correctly identified.	The section headings are correctly identified
4.	Please check whether the edit made in the sentence 'The Warm-up...short- term adaptation' is OK.	We approve the modification
5.	The term 'peg board' has been changed to 'Peg-Board' throughout the article. Please check and approve the edits made.	We approve the modification

The Validation of Surgical Simulators for RASD

Roger Smith *PhD*

Florida Hospital Nicholson Center

roger.smith@flhosp.org

http://www.nicholsoncenter.com/

FLORIDA HOSPITAL
NICHOLSON CENTER

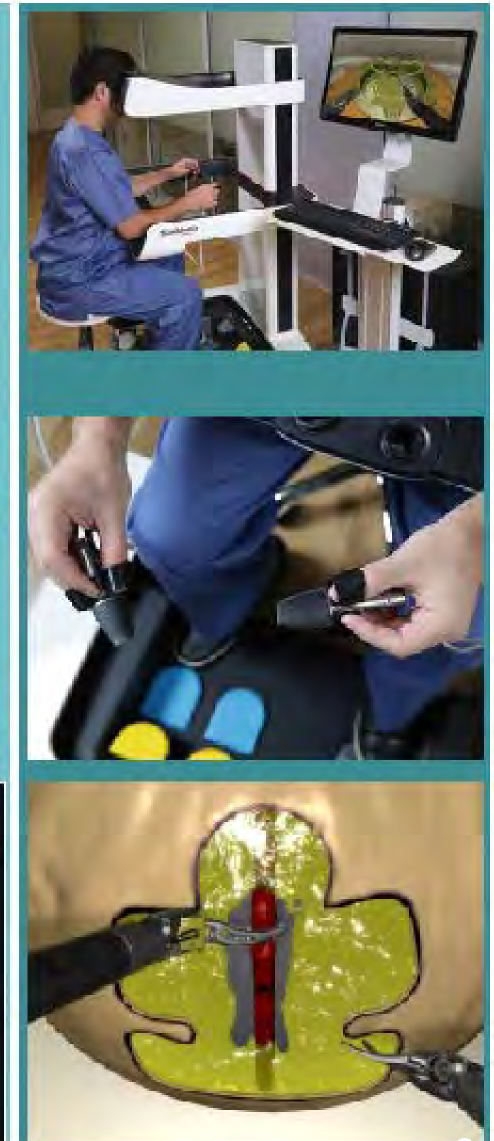
Robotic Surgery Simulators

Intuitive DVSS

Mimic dV-Trainer

Sim Surg RoSS

Simbionix Robotix Mentor



Simulator System Functionality Comparison

Features	DVSS	dV-Trainer	RoSS	Robotix Mentor
System Manufacturer	Intuitive Surgical Inc.	Mimic Technologies Inc.	Simulated Surgical Systems LLC	Simbionix Ltd.
Specifications (Complete System)	Depth 41" Height 65" Width 40" 120 or 240V power	Depth 36" Height 59" Width 54" 120 or 240V power	Depth 44" Height 77" Width 45" 120 or 240V power	Depth 36" Height 48" Width 48" 120 or 240V power
Visual Resolution	VGA 1024 x 768	VGA 1024 x 768	VGA 640 x 480	VGA 1024 x 768
Components	Customized computer attached to da Vinci surgical console	Standard computer, visual system with hand controls, foot pedals	Single integrated custom simulation device	Standard computer, visual system with hand controls, foot pedals
Support Equipment	da Vinci Si surgical console	Adjustable table, touch screen monitor, keyboard, mouse, protective cover, custom shipping container	USB adapter, keyboard, mouse	None
Exercises	35 simulation exercises (30 by Mimic, 5 by Simbionix)	65 simulation exercises	52 simulation exercises	50 simulation exercises
Optional Software	PC-based Simulation management	Mshare curriculum sharing web site	Video and Haptics-based Procedure Exercises (HoST)	Uro, Gyn Procedural Modules
Scoring Method	Scaled 0-100% with passing thresholds in multiple skill areas	Proficiency-based metric and point system with passing thresholds in multiple skill areas	Point system with passing thresholds in multiple skill areas	Proficiency-based metric and point system with passing thresholds in multiple skill areas
Student Data Management	Custom control application for external PC. Export via USB memory stick.	Export student data to delimited data file and graphical reports.	Export student data to delimited data file.	Export student data to delimited data file and graphical reports.
Curriculum Customization	None	Select any combination of exercises. Set passing thresholds and conditions.	Select specifically grouped exercises. Set passing thresholds.	Online curriculum development.
Administrator Functions	Create student accounts on external PC. Import via USB memory stick.	Create student accounts. Customize curriculum.	Create student accounts. Customize curriculum.	Create student accounts, export data, create curriculum.
System Setup	None.	Calibrate hand controls.	Calibrate hand controls.	Calibrate goggles.
System Security	Student account ID and password.	Administrator password, Student account ID and password, Guest account.	PC password, Administrator password, Student account ID and password.	Administrator password, Student account ID and password, Guest account.
Simulator Base Price	$85,000	$99,200	$126,000	$75,000
Support Equip Price	$500,000	$9,800	$0	$0
Total Functional Price	$585,000	$109,000	$126,000	$75,000

Simulator Validation Studies

Types of Validation *McDougall 2007*	**DVSS**	**dV-Trainer**	**RoSS**
Face *Subjective realism of the simulator*	*Hung 2011* *Kelly 2012* *Liss 2012* *Tanaka 2014*	*Lendvay 2008* *Kenney 2009* *Sethi 2009* *Perrenot 2011* *Korets, 2011* *Lee 2012* *Tanaka 2014*	*Seixas-Mikelus 2010* *Stegemann, 2012* *Tanaka 2014*
Content *Judgment of appropriateness as a teaching modality*	*Hung 2011* *Kelly 2012* *Liss 2012* *Tanaka 2014*	*Kenney 2009* *Sethi 2009* *Perrenot 2011* *Lee 2012* *Tanaka 2014*	*Seixas-Mikelus 2010* *Colaco, 2012* *Tanaka 2014*
Construct *Able to distinguish experienced from inexperienced surgeon*	*Hung 2011* *Kelly 2012* *Liss 2012* *Finnegan 2012* *Tanaka 2014*	*Kenney 2009* *Korets, 2011* *Perrenot 2011* *Lee 2012* *Tanaka 2014*	*Raza, 2013* *Tanaka 2014*
Concurrent *Extent to which simulator correlates with "gold standard"*	*Hung 2012*	*Lerner 2010* *Perrenot 2011* *Korets 2011* *Lee 2012*	*Chowriappa, 2013*
Predictive *Extent to which simulator predicts future performance*	*Hung 2012* *Culligan 2014*		

Robotic Skills vs. Simulator Exercises/Metrics

Core Skills	Content	Simulators			
		dV-Trainer	RoSS	DVSS	Robotix Mentor
Wristed Bimanual Manipulation	Tasks	Yes	Yes	Yes	Yes
	Metrics	No	No	No	No
Maneuver Camera	Tasks	Yes	Yes	Yes	Yes
	Metrics	No	No	No	No
Master Cntrl Clutching	Tasks	No	Yes	Yes	Yes
	Metrics	M Wrk Rng	No	M Wrk Rng	Yes
Use 3rd Arm	Tasks	Yes	Yes	Yes	Yes
	Metrics	No	No	No	Yes
Depth/3D Perception	Tasks	No	No	No	No
	Metrics	No	No	No	No
Aware Instrument Force	Tasks	No	No	No	Yes
	Metrics	Ex Inst Frc	No	Ex Inst Frc	Yes

Liu & Curet, 2015

Nature of Validation in Surgical Simulation

"Types of Validation"

AERA/APA, 1974 McDougall, 2007
Face Content Construct Concurrent Predictive

↑

As of 2010, 100% of the studies in the published literature have used this model of validation. (Korndorffer, 2010)

"Sources of Evidence"

AERA/APA, 1990 Korndorffer, 2010
Content Process Response Internal Structure Relations to Other Variables Consequences

↑

Surgical education community is still struggling to understand how to apply this model. It is not in general use at this time.

Training & Simulation to Mitigate Risk

1. **Robot.** Accessibility given cost of device (~$1.6 million) requires support from manufacturer.
2. **Simulator.** Minimal device cost, operational support costs (staff and facilities) are difficult.

3. **Curriculum.** Authoritative, standardized, objective.
4. **Metrics.** Validated, applicable across devices and curriculum.
5. **Certification.** Enforced by boards, societies, and hospitals.

Fundamentals of Robotic Surgery

addresses #3, 4, and 5.

References

Smith R, Truong M, Perez M. (Jan 2015). Comparative analysis of he functionality of simulators of the da Vinci surgical robot. Surg Endosc, 29:972-983.

McDougall EM. (2007). Validation of surgical simulators. J Endourol, 21(3):244-7.

AERA/APA. (1990 & 1974). The Standards for educational and psychological testing. http://www.apa.org/science/programs/testing/standards.aspx

Korndorffer JR, Kasten SJ, Downing SM. (2010). A call for the utilization of consensus standards in the surgical education literature. Amer J Surg, 199, 99-104.

Hung AJ, Zehnder P, Patil MB. (2011). Face, content and construct validity of a novel robotic surgery simulator. J Urology, 186(3):1019-24.

Kelly DC, Margules AC, Kundavaram CR. (2012). Face, content, and construct validation of the da Vinci Skills Simulator. J Urology, 79(5):1068-72.

Liss MA, Abdelshehid C, Quach S. (2012). Validation, Correlation, and Comparison of the da Vinci Trainer(™) and the da Vinci Surgical Skills Simulator(™) Using the Mimic(™) Software for Urologic Robotic Surgical Education. J Endourol, 26(12):1629-34.

Lendvay, P. Casale, R. Sweet, C. Peters. (2008). Initial validation of a virtual-reality robotic simulator. J of Rob Surg, 2(3):145-149.

Kenney P, Wszolek M, Gould J, Libertino J, Moinzadeh A. (2009). Face, Content, and Construct Validity of dV-Trainer, a Novel Virtual Reality Simulator for Robotic Surgery. J Urology, 73(6):1288-92.

Sethi AS, Peine WJ, Mohammadi Y. (2009). Validation of a novel virtual reality robotic simulator. J Endourol, 23(3):503-8.

Perrenot C, Perez M, Tran N. (2012) The virtual reality simulator dV-Trainer(®) is a valid assessment tool for robotic surgical skills. J Surg Endosc, 26(9):2587-93.

Korets R, Mues AC, Graversen JA, Gupta M, Benson MC, Cooper KL, Landman J, Badani KK (2011) Validating the Use of the Mimic dV-Trainer for Robotic Surgery Skill Acquisition Among Urology Residents. J Urology, 78(6):1326-30.

Lee JY, Mucksavage P, Kerbl DC. (2012) Validation study of a virtual reality robotic simulator--role as an assessment tool? J Urology, 187(3):998-1002.

Seixas-Mikelus SA, Kesavadas T, Srimathveeravalli G. (2010). Face validation of a novel robotic surgical simulator. J Urology, 76(2):357-60.

Stegemann AP, Kesavadas T, Rehman S, Sharif M, Rao A, DuPont N, Shi Y, Wilding,G, Hassett J, and Guru K. (2012). Development, Implementation and Validation of a Simulation-Based Curriculum for Robot-Assisted Surgery. Presented at the AUA poster session, May 19-23, 2012 Atlanta, GA.

Colaco M, Balica A, Su D. (2012) Initial experiences with RoSS surgical simulator in residency training: a validity and model analysis, J Rob Surg, 7(1):71-75

Finnegan KT, Meraney AM, Staff I. (2012). da Vinci Skills Simulator construct validation study: correlation of prior robotic experience with overall score and time score simulator performance. J Urology, 80(2):330-5.

Raza SJ, Chowriappa A, Stegemann A, Ahmed K, Shi Y, Wilding G, Kaouk J, Peabody J, Menon M, Kesavadas T, Guru KA. (2013). Construct Validation of the Fundamental Skills of Robotic Surgery (FSRS) Curriculum. Podium presentation, AUA Annual Congress, San Diego,

Hung AJ, Patil MB, Zehnder P. (2012). Concurrent and predictive validation of a novel robotic surgery simulator: a prospective, randomized study. J Urology, 187(2):630-7.

Lerner M, Ayalew M, Peine W, Sundaram C. (2010). Does Training on a Virtual Reality Robotic Simulator Improve Performance on the da Vinci® Surgical System? J Endourol, 24(3):467-72.

Chowriappa A, Raza SJ, Stegemann A, Ahmed K, Shi Y, Wilding G, Kaouk J, Peabody JO, Menon M, Kesavadas T, Guru KA. (2013). Development & Validation of a Composite Grading & Scoring System for Robot-assisted Surgical Training – The Robotic Skills Assessment Scale. Podium presentation, AUA Annual Congress, San Diego.

Stegemann A, Kesavadas T, Rehman S, Sharif M, Rao M, duPont N, Shi Y, Wilding G, Hassett J, and Guru K. (2013). Development, Implementation, and Validation of a Simulation-Based Curriculum for Robot-Assisted Surgery. J Urology, 81(4):767-74.

Culligan PC et al. (Jan 2014). Predictive validity of the training protocol using a robotic surgery simulator. Fem Pelvic Med & Recon Surg, 20(1): 48-51.

Liu M & Curet M. (Jan 2015). A review of training research and virtual reality simulators for the da Vinci surgical system. *Teaching and Learning in Medicine*, 27(1):12-26.

Panel: Simulation as a Means of Ensuring Device Usability and Risk Mitigation

Roger Smith *PhD*

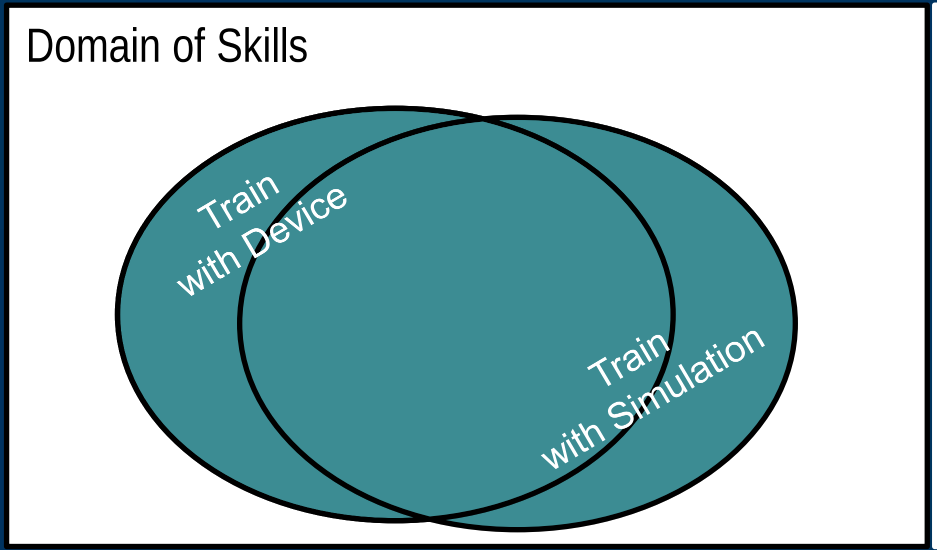
Florida Hospital Nicholson Center

roger.smith@flhosp.org

http://www.nicholsoncenter.com/

FLORIDA HOSPITAL
NICHOLSON CENTER

The Role of Simulation in Training



Advantages
Cost of device
Training support costs
Programmable curriculum
Automated metrics
Progression of skills

Disadvantages
Accuracy of replication
Completeness of replication
Additional system support
Lag device upgrades
Community acceptance

Challenge V.1: To what extent can training be provided to mitigate device specific risk?

- Obvious: Training in failure modes and recovery, safe set-up and verification, surgeon skills, team roles.
- Forms: training with equipment, dry labs, tissue, role playing, simulated emergency.
- Issues: Cost of equipment, staff, and services. Training ROI is not good enough to motivate any 3rd party to offer training. Financially only makes sense for device manufacturer.

Challenge V.2: To what extent can professional societies/hospitals address medical device training as part of curriculum?

- Create standard materials
- Establish benchmarks
- Require test & evaluation
- Issue: How do societies/hospitals develop expertise to establish these in the early phases? Bootstrap needed to get started. Need trusted partnership with device companies.

Challenge V.3: How can individual surgeons know that they have optimized their skills before treating patients?

- Cross correlate multiple training methods & curricula. Use 2-3 different methods to measure the same skill. Matching errors in each method indicate a real lack of skill, not just an isolated mistake (Polya, 1945).
- Many more tricks in the mathematic and psychometric toolkit.

Challenge V.4: What role can surgical simulation play to facilitate RASD training?

- Dry lab tools, curriculum & standards
- Cadaver and animal labs
- VR/3D sims for skills and procedures
  - Early wireframe approaches used by NASA and DOD
- Team role playing and curriculum tools
- OR staff training
- Emergency procedures course

Robotic training with porcine models induces less workload than virtual reality robotic simulators for urology resident trainees

VLADIMIR MOURAVIEV[1,] MARTINA KLEIN[2], ERIC SCHOMMER[3], DAVID D. THIEL[3], SRINIVAS SAMAVEDI[1], ANUP GUPTA[1], RAYMOND J. LEVEILLEE[4], RAJU THOMAS[5], JULIO POW SANG[6] , LI-MING SU [7] , ENGY MUI[8], ROGER SMITH[9]AND VIPUL PATEL[1]

[1]GLOBAL ROBOTICS INSTITUTE, CELEBRATION, FL; [2]UNIVERSITY OF TEXAS, LUBBOCK, TX; [3] DEPARTMENT OF UROLOGY, MAYO CLINIC, JACKSONVILLE, FL; DEPARTMENT OF UROLOGY, UNIVERSITY OF MIAMI, FLORIDA; [5] DEPARTMENT OF UROLOGY, TULANE, LA; [6] DEPARTMENT OF UROLOGICAL ONCOLOGY, MOFFITT CANCER CENTER, TAMPA , FL, [7]DEPARTMENT OF UROLOGY, UNIVERSITY OF FLORIDA, GAINESVILLE, FL; [8]UNIVERSITY OF CENTRAL FLORIDA COLLEGE OF MEDICINE, ORLANDO, FL; [9]NICHOLSON CENTER, FLORIDA HOSPITAL, CELEBRATION, FL

INTRODUCTION

Robotic assisted urologic surgery is predicted to continue to grow in usage in the coming years, and residents trained in urology will increasingly be expected to be proficient in robotic surgery[1]. The complexity of robotic technology, its steep learning curve, and work-hour limitation of resident trainees make incorporating robotic training into residency a challenging task. Experts suggest that learning as a bedside assistant for robotic surgery has a rapid plateau; many programs are now utilizing physician's assistants and surgical technicians for bedside duties in order to free the residents for console training[2]. In high volume programs it remains difficult for residents to gain hands-on console time due to their insufficient skill set and the complexity of most procedures.

Robotic simulation training tools can therefore be utilized by novice trainees to shorten the learning curve and improve operative skills in a low-risk environment. In pursuit of improving the quality of residents' education, the Southeastern Section of the American Urological Association (SES AUA) hosts an annual robotic training course for its residents. The aim of this study is to evaluate robotic simulation workload and stress levels on urology resident trainees utilizing porcine models and virtual reality robotic simulators during this workshop.

Material and Methods.

Select residents from each of the 14 training programs of SES AUA are invited to Orlando, FL, for a 2-day robotics training course. Up to3 residents were invited from each training program. The 2015 cohort of residents represented a wider range of training and diversity in experience than in previous courses being exposed to robotic surgery early at their home institutions. Volunteer faculty were recruited from multiple SES AUA training programs.

SES AUA resident assessment of the workload associated with robotic simulation and live robotic surgery in a porcine model were assessed using the NASA Task Load Index (NASA-TLX). The NASA-TLX assesses workload along six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration[3]. Each is measured on a 21-point scale between "Very Low" and "Very High" (Appendix 2).

The SESAUA robotics course undertaken by the residents is outlined below[4].

ROBOTIC COURSE DAY 1

A full didactic session broken into 3 components. Component 1 covered the basics of robotic surgery including room set-up, bedside assistance, and console essentials. Component 2 covered several aspects of robotic kidney surgery including patient positioning, port placement, and surgical techniques. Component 3 focused on robotic prostate surgery including port placement and different surgical techniques. Didactics were supplemented with surgical videos and discussions of difficult surgical scenarios and possible complications.

ROBOTIC COURSE DAY 2

The trainees were divided into two groups. Half were asked to perform skill tasks on the Mimic da Vinci-Trainer (MdVT, Mimic Technologies, Inc., Seattle, WA, U.S.A) for four hours while the other half performed set tasks in a live nephrectomy on porcine model using the da Vinci Xi robot (Intuitive Surgical Inc., Sunnyvale, CA). After the four hours the groups changed places and continued for another 4-hour session.

SIMULATION SECTION

In the 4-hour MdVT simulation session, trainees were first given a tutorial of the console and its functionality. The trainees then proceeded to complete 5 applications of increasing difficulty and required skills. The first application, “pick and place”, involved simple movements of pyramidal jacks into corresponding colored bowls and is used to orient the trainee to the simulator. The second application, “peg board” is more advanced and required the trainee to clutch hand instruments while moving the camera, which involves coordinated hand and foot movements. The third application, “ring walk”, involved moving a ring over a curved bar without touching the bar with any portion of the ring. This drill requires all the above skills as well as maintaining awareness and accuracy with the ring position in three dimensions. The fourth application, “thread the rings”, involves passing a curved needle through rings positioned at different angles without touching the ring with any part of the needle. This drill teaches trainees good suturing technique. The last application, “tubes 2”, is the most challenging and realistic. This drill is designed to replicate the performance of an urethrovesical anastomosis. It utilizes all of the above skills including accuracy, coordination, and sufficient needle control.

ANIMAL TRAINING SECTION

In the 4-hour porcine model live surgery session, all trainees spent one hour performing cystostomies and cystorrhaphies. They then spent thirty minutes practicing port insertion and robot docking. Finally, for 2.5 hours, trainees conducted a bilateral nephrectomy which included artery, vein, kidney and ureter dissections and ligation.

QUESTIONNAIRE

All trainees were asked to complete a 1-page demographic questionnaire following the MdVT session (Appendix 1). They were also asked to complete the NASA TLX 1-page questionnaire following both the MdVT simulation and live animal model sessions (Appendix 2).

STATISTICAL ANALYSIS

Descriptive (range, mean, standard deviation, frequency). *t*-tests and analysis of variance using SAS (version 9.2; SAS Institute Inc., Cary, NC, U.S.A.).

RESULTS

Twenty-one residents from 14 programs in the SES AUA participated in this course. Seventeen (80.9%) had used a robotic console during an actual surgical case, while four did not. The distribution of the different levels of training among the residents is shown in Figure 1. Unlike previous years' courses when only senior or chief residents participated, this course included more junior residents. This reflects a shift toward early exposure to robotic surgery during urology training in most academic programs. The number of robotic or laparoscopic surgeries performed or assisted in by residents at different levels of training is shown in Figure 2. Trainees' satisfaction with their program robotic surgery training was assessed (Figure 3). Of the 17 residents who performed actual robotic surgery, 7 (41.2%) stated that the simulator replicates real-life robotic surgery, while 10 (58.8%) stated that it did not.

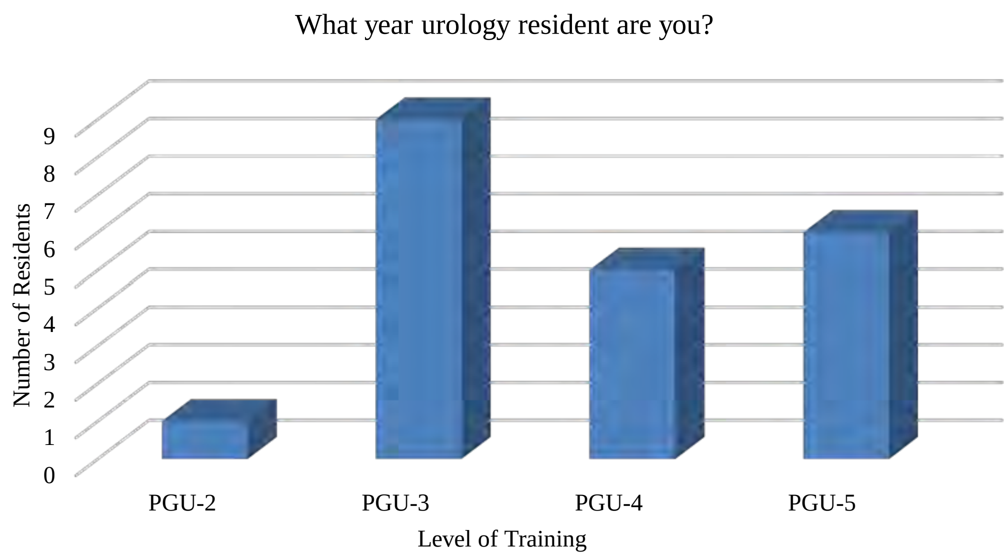


Figure 1. Robotic Simulator Questionnaire: Question 1 results

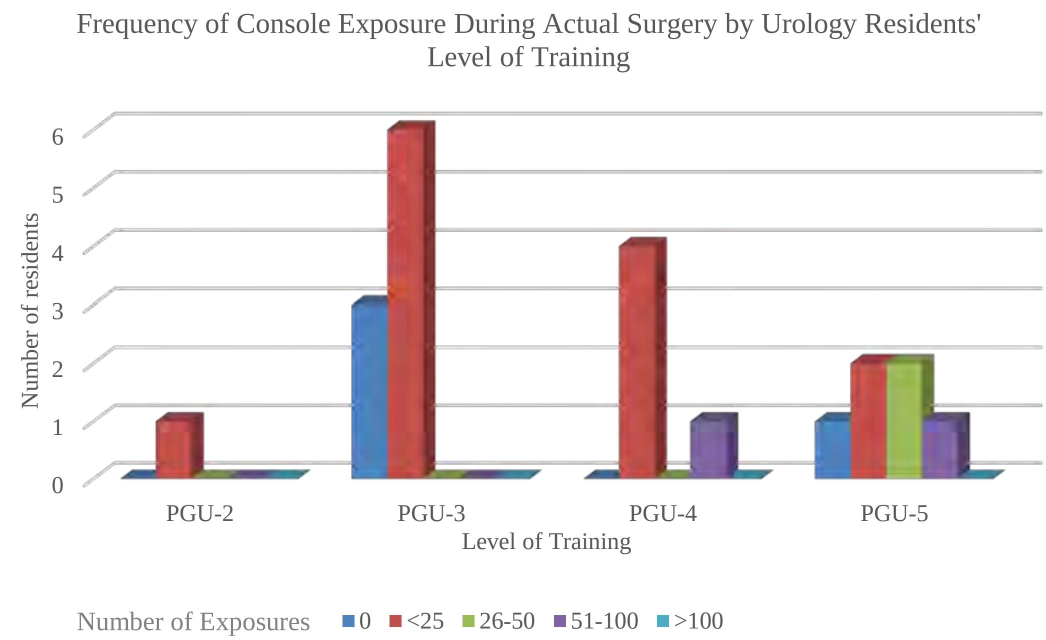


Figure 2. Robotic Simulator Questionnaire: Question 4 results

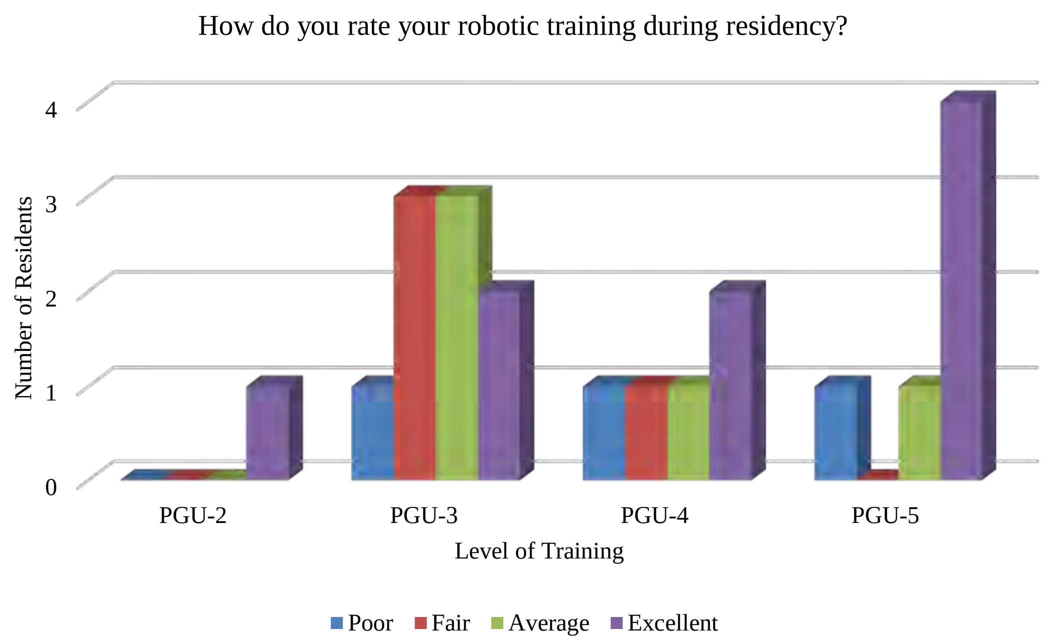


Figure 3. Robotic Simulator Questionnaire: Question 5 results

The NASA-TLX scores were converted to a 0-100 scale with 5 point increments. The raw TLX method was employed to eliminate the weight variability of the different TLX scales. To assess the NASA-TLX data at two interfaces (simulator vs. animal model) for the different levels of training (year of residency), a 4 x 2 x 6 (residency, interface, and TLX scale, respectively) mixed ANOVA was computed. The Greenhouse-Geisser correction was used to correct for the sphericity assumption. The ANOVA indicated a significant main effect for TLX scales, F (3.91, 66.44) = 4.93, $p = 0.002$, $\eta_{partial} = 0.225$, as well as a significant interface by TLX scales interaction, F (3.73, 63.42) = 3.73, p = 0.016, $\eta_{partial} = 0.166$). None of the other main effects and interactions were significant. To further analyze the TLX main effects, Bonferroni-corrected repeated-measures *t*-tests were computed to determine which TLX scales differed significantly from each other; type-I error rate per comparison was set to 0.003. Means of the TLX scales are presented in Figure 4. As can be seen in Figure 4, effort resulted in the highest score. The Bonferroni corrected t-tests indicated that mental demand was significantly higher than physical demand (t (20) = 4.05, $p = 0.001$) and then frustration (t (20) = 3.52, $p = 0.002$). Further, temporal demand was significantly higher than physical demand (t (20) = 2.90, p = 0.009) and that effort was significantly higher than physical demand (t (20) = 6.52, p <0.001)), temporal demand (t (20) = 5.12, $p < 0.001$), performance (t (20) = 5.15, $p < 0.001$), and frustration (t (20) = 6.90, $p < 0.001$).

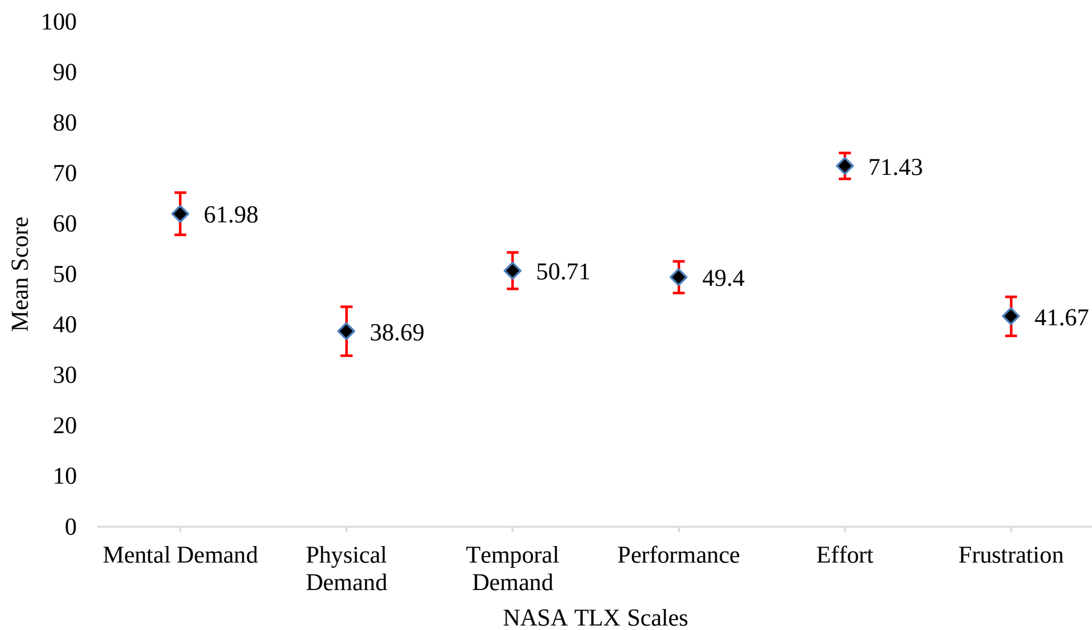


Figure 4. Mean scores of the NASA-TLX scales. Note: Error bars refer to standard error of the mean.

The analysis of the interface by TLX interaction was further analyzed to determine whether the scores of each of the six TLX scales varied across the two interfaces. On that end, Bonferroni-corrected repeated-measures *t*-tests were computed; type-I error rate per comparison was set at alpha = 0.008. The means of the TLX scores observed at the two interfaces are in Figure 5. The only significance was observed for frustration, which was significantly higher at the simulation than the animal model, $t(20) = 4.12$, $p = 0.001$).

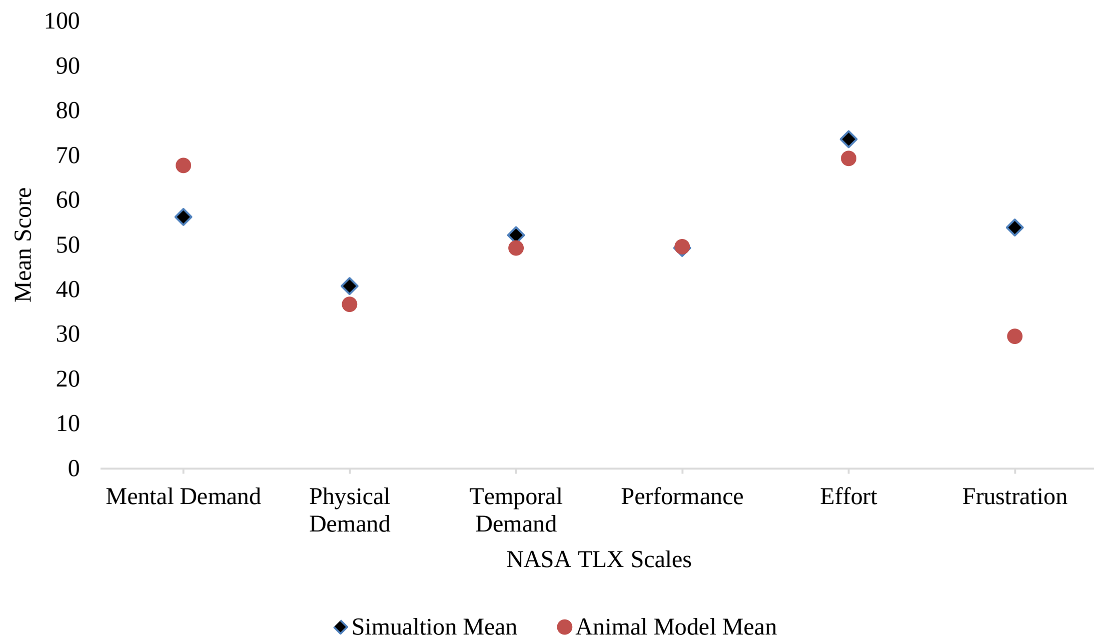


Figure 5. Mean scores of the NASA-TLX scales in simulation versus animal model

DISCUSSION

Robotic surgery is increasing in popularity in the field of urology due to its minimal invasiveness, reduced risk of complications and shortened hospital stay. This growing trend is evident in our results. The majority of the trainees in the 2015 course (80%) reported live console exposure. In contrast with a similar survey conducted in 2013 in a group of SES AUA trainees, only 56.9% of the trainees that year reported having had robotic console time[4]. During the 2014 annual training course 92% of the trainees reported performing live robotic surgery at their home institution[5]. Despite these increasing numbers, there is a lack of standardization and certification process for urology residents in robotic surgery. Furthermore, there is no standardized training protocol for residents learning robotic surgery across the various training programs. Gover et al. suggested a threshold of 25-30 cases for a novice surgeon to begin to operate the foot pedals and controls safely and intuitively[6]. Only 4 (19%) of our trainees reported having performed more than 25 cases.

Robotic surgery simulators have been proposed to narrow the gap of novice trainees' skill levels[7]. They would also help establish the basics of important operative skills such as eye-hand-foot coordination and comfort using the console controls and foot pedals. Our program chose to use the Mimic dV-Trainer simulator for training, which is one of the most established virtual robotic surgical simulators today. The current version of the dV-Trainer (version 2) contains 51 exercises organized into nine categories[7]. This device also includes video and audio instructions on how to use the robotic equipment, moving through progressively more difficult skills. Kenney et al. suggested a content and construct validity in their study[8]. There are other robotic simulators that are commercially available such as the da Vinci Skills Simulator (DVSS) (Intuitive Surgical Inc., Sunnyvale, CA, U.S.A.), the Robotic Surgical Simulator (RoSS; Simulated Surgical Systems, Buffalo, NY, U.S.A.) and Robotix Mentor™ (3D Systems, Simbionix Products, Cleveland, OH, U.S.A.). However, they do not meet yet the same degree of validity as MdVT.

Recently, Lerner et al. demonstrated that residents who trained on the MdVT outperformed those who trained solely on the real da Vinci surgical system (Intuitive Surgical Inc., Sunnyvale, CA) when taking a robotic skills assessment using the real da Vinci system[9].

However, even MdVT simulator usage is not without its limitations. The MdVT costs from $85,000 to $100,000 to purchase with additional annual maintenance fees. Additionally there are no surgical procedure simulation drills, only specific skill tasks like the ones uses in this group. There are currently no urologic-specific procedure modules or simulation drills available, only general surgical skill tasks like the ones used in this group. This limitation could limit the pace of learning and might not translate to better operative skills without supplemental live surgery console time. Therefore, work on more realistic 3D case simulations to advance clinical decision-making and procedural knowledge is currently in progress. The animal model used in this analysis costs roughly $500/hr[10]. It also lacks realistic human anatomy and might provide a false sense of security which could increase risks to future patients[11]. Future work should focus on developing urology-specific training modules such as radical prostatectomy and partial nephrectomy simulations. The existing application only hones skills used in general robotic surgery and is not necessarily reflective of skills needed to perform urologic robotic surgery.

Educators and companies have yet to determine the best model to use for teaching robotic surgery. Many factors must be taken into consideration including the cost, availability of expert faculty, legal responsibility of supervising faculty, risk to patients, and the additional workload on trainees. This analysis has been conducted with aim to assess the latter factor. Trainees agreed that effort exceeded the other five workload dimensions asked on the NASA-TLX. Trainees felt that they worked physically and mentally hard to accomplish their tasks both on the simulator and the live animal model. In general, the performance level was the same for both parts of training which suggests that the training accomplished its goal.

Mental demand greatly surpassed physical demand, as expected. The trainees also reported being less frustrated with the live animal model than with the simulator. This could be due to the trainees' familiarity with live anatomical structures over skill set simulations which remains a real challenge to novice surgeons. Simulators also provide metrics to score specific performance traits, as well as combining all of these into a single composite score of performance for the entire exercise[7]. In addition to the objective metrics the MdVT simulator defines thresholds which indicate whether the trainee's score is considered a "passing" or "failing" performance with acceptable and warning scoring levels, respectively[7]. These thresholds were derived from data collection and analysis with a large number of experienced surgeons. Therefore, higher mental demand and frustration levels of trainees with simulation may suggest a bigger challenge and effort to accomplish the task sequentially through multiple repetitions of an exercise in order to reach the desirable "passing score".

Conversely, the animal hands-on part of the course did not have objective metric parameters to assess the skill set of trainees in robotics. The faculty of the course subjectively evaluated the proficiency levels of residents. Furthermore, the timeframe for every trainee with the robotic console was limited compared to the simulation part.

These results combined with previous and future SES AUA training courses' results can significantly enhance our efforts to establish a standardized robotic surgery training program that is cost effective, practical, and of the highest quality. Encouraging the development of urology-specific robotic training tools in simulation will also aid in reaching our goal. Some limitations of this analysis include its regional focus and limited sample size. The analysis also did not assess the methods each program uses for robotic training. Upon completion of residency programs, many urologists appreciate the effort and learning curve associated with robotic surgery, and believe that training and proficiency in robotic surgery are necessary during residency[9]. Future direction for this project includes compiling detailed accounts of trainees' exposures at their home institutes following the training. Such analysis combined with ongoing performance scores and trainees' subjective opinions could lead to identifying the most effective methods of training. Work is currently in progress to improve the current robotic training methods.

CONCLUSIONS

Novice trainees experienced significant mental workload while performing tasks on both the simulator and the live animal model during the robotics course. NASA TLX scoring demonstrates that live animal models provide the same proficiency performance with less frustration. The simulation part of course remains more challenging task for trainees with more frustration and repetitive exercises to achieve the passing score.

ACKNOWLEDGEMENT

The authors would like to cordially thank members of Global Robotic Institute Ms. Ashley Fialkowski and Ms. Kim Straw for their assistance in obtaining the questionnaires from participants.

REFERENCES

1. Duchene DA, Rosso F, Clayman R, McDougall EM, Winfield HN. Current minimally invasive practice patterns among postgraduate urologists. *Journal of endourology / Endourological Society.* 2011;25(11):1797-1804.
2. Thiel DD, Lannen A, Richie E, Dove J, Gajarawala NM, Igel TC. Simulation-based training for bedside assistants can benefit experienced robotic prostatectomy teams. *Journal of endourology / Endourological Society.* 2013;27(2):230-237.
3. Hart SG. NASA- Task Load Index (NASA-TLX); 20 years later. Paper presented at: Human Factors and Ergonomics 50th Annual Meeting2006; Santa Monica, CA.
4. Thiel DD, Patel VR, Larson T, Lannen A, Leveillee RJ. Assessment of robotic simulation by trainees in residency programs of the Southeastern Section of the American Urologic Association. *Journal of surgical education.* 2013;70(5):571-577.
5. Ruparel RK, Taylor AS, Patel J, et al. Assessment of virtual reality robotic simulation performance by urology resident trainees. *Journal of surgical education.* 2014;71(3):302-308.
6. Grover S, Tan GY, Srivastava A, Leung RA, Tewari AK. Residency training program paradigms for teaching robotic surgical skills to urology residents. *Current urology reports.* 2010;11(2):87-92.
7. Smith R., Truong M. Simulation in robotic surgery. Model benders Press. 46p.
8. Kenney PA, Wszolek MF, Gould JJ, Libertino JA, Moinzadeh A. Face, content, and construct validity of dV-trainer, a novel virtual reality simulator for robotic surgery. *Urology.* 2009;73(6):1288-1292.
9. Lerner MA, Ayalew M, Peine WJ, Sundaram CP. Does training on a virtual reality robotic simulator improve performance on the da Vinci surgical system? *Journal of endourology / Endourological Society.* 2010;24(3):467-472.
10. Perrenot C, Perez M, Tran N, et al. The virtual reality simulator dV-Trainer((R)) is a valid assessment tool for robotic surgical skills. *Surgical endoscopy.* Sep 2012;26(9):2587-2593.
11. Hoffman MS. Simulation of robotic hysterectomy utilizing the porcine model. *American journal of obstetrics and gynecology.* Jun 2012;206(6):523 e521-522.

Appendix 1.

Robotic Simulator Questionnaire

1. What year urology resident are you?
   - ☐ Uro-1
   - ☐ Uro-2
   - ☐ Uro-3
   - ☐ Uro-4
   - ☐ Uro-5
2. Does your training program own or have access to a robotics simulator?
   - ☐ No
   - ☐ Mimic Simulator
   - ☐ Ross Simulator
   - ☐ Mimic Backpack or console
   - ☐ Other___________
3. Have you been on the robotics console for an actual case?
   - ☐ Yes
   - ☐ No
4. Approximate the number of cases on which you have robotics console time
   - ☐ <25
   - ☐ 26-50
   - ☐ 51-100
   - ☐ >100
5. How do you rate your robotic training during residency?
   - ☐ Poor
   - ☐ Fair
   - ☐ Average
   - ☐ Excellent
6. In your experience, do you feel that the simulator replicates real life robotics?
   - ☐ Yes
   - ☐ No
7. Which drill did you find the most difficult?
   - ☐ Peg board
   - ☐ Ring Walk
   - ☐ Thread the rings
   - ☐ Tubes 2
8. If your program lacks a robotics simulator, do you think this device would be helpful in your program?
   - ☐ Yes
   - ☐ No

Appendix 2.

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High